A Bumpy Ride along the Kuznets Curve: Consumption and Income Inequality Dynamics in Russia

May 28, 2008

Yuriy Gorodnichenko
UC Berkeley
ygorodni@econ.berkeley.edu

Klara Sabirianova Peter
Georgia State University
kpeter@gsu.edu

Dmitriy Stolyarov
University of Michigan
stolyar@umich.edu

FIRST DRAFT. COMMENTS ARE WELCOME.

Abstract

We construct key household and individual economic variables using a panel micro data set from the Russia Longitudinal Monitoring Survey (RLMS) for 1994-2005. We analyze cross-sectional income and consumption inequality and find that inequality decreased during the 2000-2005 economic recovery. The decrease appears to be driven by falling volatility of transitory income shocks. The response of consumption to permanent and transitory income shocks becomes weaker later in the sample, consistent with greater self-insurance against permanent shocks and greater smoothing of transitory shocks. Comparisons of RLMS data with official macroeconomic statistics reveal that national accounts may under-estimate the extent of unofficial economic activity, and that the official consumer price index may overstate inflation and be prone to quality bias.

Keywords: inequality, income, consumption, transition, Russia.

JEL Classification: E20, J31, I30, O15, P20

1. Introduction

Modern macroeconomists are increasingly relying on analyses of environments with heterogeneous agents. The time series behavior of population distributions of macroeconomic variables is thought to affect the properties of the business cycle and the welfare effects of policy changes. Besides, many questions of interest to macroeconomists can only be asked (and answered) in the context of multi-agent environments.

These richer macroeconomic models require a correspondingly rich set of empirical facts that come from micro data and incorporate information on distributions in addition to the usual aggregates. This paper's goal is to provide a comprehensive set of cross-sectional and time series stylized facts for the Russian economy.

With this goal in mind, the present paper constructs the key variables describing the economic behavior of Russian households and individuals and analyzes their cross-sectional dispersion and time series patterns. Specifically, we create time-varying distributions of individual earnings and labor supply, as well as household-level income, expenditure, and consumption. Our primary data source is the Russia Longitudinal Monitoring Survey; a large panel data set covering the period between 1994 and 2005.

We would like to highlight two main results. First, almost all measures of cross-sectional inequality in income and consumption started falling during 2000-2005, after staying relatively high during 1994-1998. Second, the measured fall in inequality is mostly due to the moderation of transitory shocks to household income and consumption.

The recent period of falling inequality was preceded by an initial rise in the early 1990s that accompanied Russia's transition from a centrally planned to market economy (e.g., Commander *et al* 1999, Galbraith *et al* 2004). However, the level of inequality at the end of our sample is still higher than it was during the pre-transition era. Thus, the time pattern of inequality in Russia during the transition is the inverted U-shape reminiscent of the Kuznets curve.[1] Importantly, poor households do not appear to fall behind during the economic recovery

[1] Kuznets (1955) original paper explained the curve in terms of the transition between agriculture and industry. Likewise, the transition from the inefficient state sector to the new private sector may generate a similar curve. What is different in Russia and many other transition economies is that the initial rise in inequality was rapid, which

– the lower tail of the expenditure distribution does not diverge from the middle as the economy expands.

The latest level of inequality that we find is typical for a middle income country. For example, the Gini coefficient in 2005 was about 0.38-0.40, which is just slightly above the mean value of Gini coefficients for after-tax household income and consumption from upper middle income countries (WIDER v2.b, authors' calculations).[2] Some features that set the Russian economy apart from more developed countries turn out to be important for the analysis of inequality. One such feature is home production of food. Our results indicate that home-grown food has a large equalizing effect on consumption. The effect is large, because poorer rural households are also the ones that grow a lot of food for own consumption. Another unique feature of the Russian economy is its geographic diversity. Accounting for regional differences in the cost of living (that vary by a factor of 2.7 in Russia) is shown to have a sizeable equalizing effect. We argue that wage payment delays, irregularities in government transfer payments, forced in-kind substitutes in lieu of wage payments, large food inventories and other peculiar phenomena of the Russian transition in the 1990s have a significant impact on inequality measures. At the same time, we find that the variability in working hours, net private transfers, and capital income only marginally alters the level of income inequality.

We look at inequality dynamics between groups in our sample. We find the comparisons of economic experience between the urban and the rural populations particularly interesting. The rural population has a more restricted choice of jobs, which limits occupational mobility during transition. The workers with highest earnings potential might have migrated to cities, generating selection. Despite all this, we do not find evidence that the rural population fell behind. The

can be attributed to mis-pricing of labor in the Soviet system, big bang liberalization, fast mass privatization, and the failure of government redistribution policies during the first years of transition.

[2] Our results on inequality levels have to be taken in the context of our sample. We think that the RLMS, like most household surveys, may under-represent the very rich individuals who own capital assets in Russia. This is evident from the negligible financial asset holdings of most RLMS respondents. The studies that attempt to adjust for super-rich typically document much higher levels of inequality. For example, Guriev and Rachinsky (2006) find that the income Gini coefficient for the city of Moscow is 0.625, and Aivazian and Kolenikov (2001) report a Gini coefficient of 0.55-0.57 based on parametric estimation of the uncensored expenditure distribution. We find some evidence that suggests divergence between the super-rich and the rest of the population in 2003-2005 (see Section 2 for further discussion).

rural group did not seem to do relatively worse during the downturn, although during the recovery the rural population exhibited a slower growth rate in consumption of durables.

More broadly, we have found almost no evidence of convergence or divergence between groups based on observables, such as education, location, household composition and age. In other words, differences in income and consumption growth across households are not forecastable. The reduction of inequality observed during economic recovery resulted mostly from the moderation in the residual volatility of income and consumption growth.

We examine the reasons for the observed fall in residual income volatility by exploiting the panel dimensions of the data. In particular, we identify permanent and transitory shocks to income and estimate the effect of these shocks on consumption. It turns out that the fall in residual income volatility is mostly due to a fall in the variance of transitory income shocks.[3] Over time, consumption response to both permanent and transitory income shocks becomes weaker. This is consistent with better insurance against income shocks and hence better consumption smoothing later in the mid 2000s.

Apart from the analysis of inequality trends, we make levels comparisons between RLMS statistics and the macroeconomic aggregates from the National Income and Product Accounts (NIPA), which reveal some interesting results. We find income underreporting in the RLMS. At the same time, per capita expenditure levels in NIPA and RLMS match fairly well, except for the period after 2003 when the two diverged. One would expect RLMS expenditures to be lower, because this data set does not include consumption by the super-rich. The fact that RLMS expenditure matches (and sometimes exceeds) official measures probably implies that the official statistics make an insufficient adjustment for informal economic activity.

The time trends show a 40 percent drop in real per-capita expenditure and a 50 percent drop in real hourly wages during 1994-1998. Recent literature has argued that the drop in Russian real output during the transition has been overstated due to exaggeration of the Soviet output and mismeasurement of the unofficial economy in the 1990s (Schleifer and Treisman 2005) or due to overstatement of inflation by the official CPI (Gibson *et al* 2004).

[3] Stillman (2001) finds that RLMS expenditures respond strongly to transitory shocks during 1994-1998.

To detect possible sources of CPI bias, we examine food prices and quantities from RLMS and find that the composition of food consumption has not changed much. Consequently, the CPI substitution bias within the food category is likely to be small. We do find, however, substantial disagreement in food CPI between RLMS and NIPA, with a 25 percent discrepancy in the cumulative inflation during 1994-1998, but not much discrepancy afterwards. In addition, there is evidence of quality bias in the official CPI. The budget share of food kept falling throughout the downturn and economic recovery, indicating a fall in the relative price of non-food goods. Since the NIPA deflators do not show a fall in the relative price of goods, we suspect that some of the quality improvements in non-food goods may have been counted as inflation, generating a quality bias.

The paper's goal of documenting a comprehensive set of macroeconomic facts for Russia links it to many bodies of literature in macroeconomics, labor economics, development economics, and transition economics. In the interest of space, the literature survey below is necessarily incomplete, and it merely catalogues some of the related recent work by topic. Our analysis is closely related to the growing empirical literature that analyzes the joint evolution of income and consumption distributions (Cutler and Katz 1992, Attanasio and Davis 1996, Blundell and Preston 1998, Slesnick 2001, Krueger and Perry 2006, Heathcote *et al* 2007, Blundell *et al* 2008, etc.). There is also a growing body of research on inequality in developing countries. We find this literature particularly relevant for our study as it emphasizes the importance of measurement issues, urban-rural differences, home production, and income underreporting in understanding inequality in developing countries (e.g., Chen and Ravallion 1996, Deaton 1997).

Several papers document changes in income inequality in Russia in the 1990s. These studies establish a number of important facts for the early transition period: rising income inequality, significant income mobility, large regional variation, and insufficient government transfers to offset an increase in wage inequality (Commander, Tolstopiatenko and Yemtsov 1999; Milanovic 1999; Flemming and Micklewright 2000). The rise in income inequality is mainly attributed to compositional shifts from the old state sector to the new private sector, liberalization of wage setting, liberalization of prices and trade, and macroeconomic volatility. Some studies argue in favor of inequality measures based on expenditures (Aivazyan and Kolennikov 1999, Jovanovic 2001). They find a significant share of the transitory component in shocks to expenditures, high instability, and a slight downward trend in expenditure-based inequality. Our findings are in general agreement with these studies. We extend previous analyses in a number of ways. We consider a longer time span covering recent years, provide a variety of measures and decompositions of inequality, investigate sources of inequality, and examine the co-movements between income and consumption using the panel aspects of the data.

The rest of the paper is organized as follows. In Section 2, we describe the data, provide basic information on the levels of consumption, income and labor market participation, and compare these statistics with official data. In Section 3, we document the trends in inequality in individual labor market outcomes over 1994-2005. In Section 4, we construct and report consistent time series for a variety of measures of consumption and income inequality at the household level. Section 5 decomposes the income shocks into transitory and permanent components and investigates the interaction of consumption and income inequality at the household level. In Section 6, we examine the role of regional disparities and dispersion of prices in generating inequality and discuss the possible sources of CPI bias. Our concluding remarks are in Section 7.

2. Data Overview

Sample and variables

The analysis in this paper uses the Russian Longitudinal Monitoring Survey (RLMS), which is a panel dataset that includes detailed information on measures of income, consumption, household demographics, and labor supply. RLMS is organized by the Population Center at the University of North Carolina in cooperation with the Russian Academy of Sociology. The data are collected annually, and our panel includes 10 waves during the period 1994-2005, with the exception of 1997 and 1999, when the survey was not administered.[4] There were approximately 8,343-10,670 individuals who completed the adult (age 14 and over) questionnaire and 3,750-4,718 households who completed the household questionnaire in each round. These individuals and households reside in 32 oblasts (regions) and 7 federal districts of the Russian Federation.[5]

The RLMS sample is a multi-stage probability sample of dwellings. The response rate is relatively high: it exceeds 80% for households and about 97 percent for individuals within the households. The sample attrition is generally low compared to similar panel surveys in other countries, partly owing to lower mobility and infrequent changes of residences.[6] To account for the panel attrition, all statistics reported in this study are weighted using the RLMS sample weights that adjust not only for sample design factors but also for deviations from the census characteristics. For comparability with other countries in this volume, we restrict our estimation sample to households in which at least one individual is 25-60 years old. Appendix 3 shows the size and composition of the estimation sample.

The variables employed in our study are carefully constructed and made not only internally comparable across different waves but also externally consistent with standard variable definitions in macroeconomic literature. We provide thorough treatment of missing values, influential observations, non-response, and other common problems of micro data. We also take into account important Russia-specific phenomena that influence our variable definition and data analysis such as wage payment delays in the 1990s, production of food at home, high regional diversity in cost of living, as well as peculiarities of the transition to a market economy. The detailed procedures of variable construction are documented in Appendix 1.

[4] In all plots except Figure 2, the 1997 and 1999 values are 2 point linear interpolations of the data points in adjacent years.

[5] Russia had 89 regions and 7 federal districts as of December 1, 2005. The RLMS sample consists of 38 randomly selected primary sample units (municipalities) that are representative of the whole country.

[6] To deal with attrition, RLMS replenishes its sample on a regular basis by adding new dwellings, especially in the areas of high mobility such as Moscow and other large cities. To maintain the panel, RLMS partially attempts to collect information on those who moved out of the sample dwellings but live in the same location. More details on sample design, attrition, and replenishment are available at http://www.cpc.unc.edu/projects/rlms.

Economic conditions

Economic conditions in Russia affect our interpretation of income and consumption data in important ways. During the 1994-2005 period, Russia continued its transformation from a centrally planned system into a market economy. New integrated markets have emerged and new institutions of private ownership and property rights have been established.

This transition to a market economy was accompanied by extreme macroeconomic disturbances, both real and nominal. Our sample period features two distinct phases: the downturn in 1994-1998 and the post-1998 period of rapid recovery. Panel A of Figure 1 shows that the early 1990s, following price liberalization in 1992, was a period of hyper-inflation. The end-year inflation rate in 1994 was 214 percent. The 1998 inflation spike (84 percent) corresponds to the government default on sovereign debt and the abrupt devaluation of the national currency, the ruble. In the downturn, real per-capita income and expenditures fell by about 40 percent (see panels B-D). Employee compensation and public transfers were paid irregularly, and were delayed by 3 to 5 months, on average. In the recovery, real per-capita income and expenditure growth was around 9 percent annually, and inflation stayed relatively low (10 to 20 percent).

Composition of income

The composition of household income during the sample period remained relatively stable, although there are important differences with Western industrialized economies. Panel B of Figure 1 compares different measures of household after-tax monthly income during 1994-2005. Labor income, yL, is by far the largest income source; it accounts for 82 percent of household after-tax disposable income, yD, on average. Income derived from financial assets is negligible; there is only a tiny difference between y and $yL+$. The impact of net private transfers on disposable income is also small, as can be seen from the gap between $yL+$ and yL. These are contributions in money and in kind received from friends, relatives, and charitable organizations minus contributions given to individuals outside the household unit. Although net private transfers are small, gross private transfers are significant: the amount of private transfers *received* is about 9 percent of *yD*, which is comparable to the amount of public transfers. The share of public transfers is 13 percent on average, and it has increased after 2001, as evidenced by the growing gap between *yD* and *y* in panel B.

Composition of expenditures

Household consumption is constructed from numerous disaggregated categories of expenditures. Non-durable items, *c*, include 50 categories of food at home and away from home, alcoholic and non-alcoholic beverages, tobacco products, expenses on clothing and footwear, gasoline and other fuel expenses, rents and utilities, and 15-20 subcategories of services such as transportation, repair, health care services, education, entertainment, recreation, insurance, etc. Durable consumption is based on purchases of durable items within the last 3 months. All consumption measures are converted to a monthly base. To keep the coverage of consumption consistent across years, we exclude expenditure categories that only became available in recent years, such as washing supplies, personal hygiene items, books, sporting equipment, internet, and wireless phone services.

Food is the biggest expenditure category for most households. The share of food purchases in aggregate non-durable expenditures starts from a high of nearly 70 percent in 1994 and gradually falls to 49 percent in 2005 (see also Figure 1C). One peculiar feature of Russian households is that many of them grow agricultural products on their subsidiary plots for own consumption. Thus, food grown at home is an important part of consumption. In 1994, about 10 percent of total food consumption (by market value) was home-grown, and by 2005 the share of food grown at home declined to 5 percent (see also Figure 1C). Home production of food is concentrated among the poorer rural households, and although its aggregate importance is declining, it has significant economic effects on measures of inequality.

The share of durables was around 14 percent of aggregate expenditures, *cD*, during 1994-2002, but has increased significantly after 2003 (see also Figure 1D). Expenditures on durables tend to be concentrated at high income levels. 76 percent of households report no durable purchases within the last 3 months.

For 2000-2005, our dataset has a self-reported market value of owner-occupied housing. If we take the annual housing services flow to be 5 percent of its market value, the share of owner-occupied housing will equal roughly 11 percent of total consumption, *cD+*. The share of housing consumption is relatively stable over time because the aggregate market value of housing is growing at roughly the same rate as aggregate expenditures, *cD* (see Figure 1D).

Income under-reporting

Panel D of Figure 1 compares the levels of income and consumption in our sample. Expenditures are consistently above income throughout the whole period. This is not attributed to dissaving, as most households have negligible stocks of financial assets. We believe that income is under-reported because of tax evasion. For example, Gorodnichenko *et al* (2008) studied the gap between consumption and income in the RLMS data set and found the gap to be significantly larger in districts where respondents believed that other people do not pay their taxes. Over time, the gap between consumption and income seems to narrow, and the narrower gap may correspond to the effect of the 2001 tax reform (see Gorodnichenko *et al* 2008).

Since we do not have an independent estimate of the extent of income under-reporting (except the consumption-income gap itself), it seems to be more informative to compare expenditure, rather than income, levels between RLMS and other data sources.

Comparison with national accounts

We first compare income and expenditure levels between RLMS and official National Income and Products Accounts (NIPA). To make comparisons with national statistics, one must be careful about using compatible data definitions. The RLMS measure of household disposable income (*yD*) is after taxes and transfers *given*, and it excludes in-kind consumption, such as owner-occupied housing and home-grown food. The matching NIPA measure is disposable income for the "household account" after taxes and transfers minus in-kind consumption (Goskomstat 2007a). Similarly, the RLMS measure of consumption that we select for comparison purposes (*cD*) corresponds to the NIPA measure of household final consumption expenditures on durable and non-durable goods and services without imputed in-kind expenditures (Goskomstat 2007a). For comparability purposes, we use the full unrestricted sample.

Panels A and B of Figure 2 compare *yD* and *cD* (in per capita terms) with their counterparts from NIPA. Consumer expenditures in RLMS and NIPA are close during most of the sample period[7], while reported disposable income in RLMS is up to 30 percent lower than the official figures. The big discrepancy in income levels across the two sources is expected, since NIPA expenditure and income data are internally consistent,[8] and RLMS reported income is much lower than expenditures. This comparison supports income under-reporting as a possible explanation for the consumption-income gap in the RLMS and also points to expenditure data as potentially more informative about the level variables.

The close agreement between RLMS and NIPA expenditure numbers in panel B is somewhat more surprising. This finding contrasts sharply with similar comparisons in the U.S. literature that finds that household surveys tend to underestimate national aggregates by more than 30 percent. The analogous comparisons for the UK produce a less significant discrepancy of 5 percent (Attanasio *et al* 2004). One would expect expenditures to be higher in NIPA since household surveys are known to under-represent the very rich individuals. The under-representation of the extremely wealthy in the RLMS is implicitly confirmed by the very small capital income among the RLMS households. Since RLMS aggregate capital income is negligible, the NIPA expenditure should, in theory, exceed the RLMS expenditure by the amount equal to aggregate consumption from capital income. The capital owners possess significant wealth: for example, Guriev and Rachinsky (2006) estimate that the personal wealth of Russian billionaires and millionaires equals 1.4 times the country's GDP. Even if this group earns a conservative return on their wealth, they should account for at least 10-12 percent of aggregate expenditures. The fact that NIPA expenditures in panel B is not much higher than its RLMS

[7] The 1998 discrepancy can be explained by the fact that RLMS has been conducted just right after the August financial crisis while NIPA's numbers are averaged over the year.

[8] NIPA eliminates the discrepancy between reported income and consumption by construction. Disposable income is constructed as a sum of household aggregate expenditures and savings, and the difference between imputed disposable income and the officially reported income is included in the income accounts as unobserved labor compensation.

counterpart probably implies that the official statistics make an insufficient adjustment for shadow economic activity, or that they under-estimate the resources of the super-rich.

Starting in 2003, RLMS consumption expenditures show slower growth than NIPA expenditures. As explained above, this difference in trends may indicate the growing gap between the RLMS sample and the super-rich individuals. Part of the gap may also be due to an upward trend in consumption of goods that RLMS data does not consistently track, such as internet and cell phone services. However, new consumption categories added to RLMS over the years account for at most 8 percent of aggregate expenditures, and their growth is not enough to account for the difference in trends after 2003. Finally, a small portion of the gap (up to 1.6 percent of aggregate expenditures per capita) can be explained by the replacement of one of the wealthiest oil-based regions in the North by the middle income region in Siberia in the 2003 RLMS sample (this was the only episode of regional sample replacement during the 1994-2005 period).

Comparison with the Household Budget Survey

We also compare RLMS with another official data source, the Household Budget Survey (HBS). HBS is the core Goskomstat source for published statistics on income differentiation and the composition of income and consumption. The HBS micro files are not publicly available. It is worthy of note that Goskomstat does not publish the actual income levels from HBS possibly for the reasons of massive under-reporting. Instead, it imputes money income as the sum of household expenditures and changes in financial assets (Goskomstat 1999).

Panels C and D of Figure 2 show the trends in consumption of food (including food grown at home) and non-food items, respectively. The statistics reported in Panels C and D are the average monthly consumption expenditures per household member. The RLMS expenditures are about 20 percent higher than its HBS counterpart, with the discrepancy being larger for non-food items. Some of the discrepancy is due to RLMS survey timing: the HBS reports average monthly consumption in a given year while RLMS reports last month consumption at the end of year. Then RLMS numbers should be larger when there is an upward trend in consumption and smaller when there is a downward trend. It is also plausible that consumer expenditures in HBS are underreported if people are reluctant to reveal their actual level of well-being in an official survey that asks, among other expenditures, the amount of taxes paid (RLMS does not ask about taxes, nor it is linked to any government agency).

Overall, RLMS appears to be a reliable data source for examining the inequality trends in labor market outcomes, reported income, consumption, with the common caveats of income under-reporting and underrepresentation of the super-rich.

3. Inequality in Labor Market Outcomes

Since labor income is the most prevalent income source, the inequality in labor market outcomes is crucial for understanding the overall income inequality. This section takes a closer look at the dynamics of inequality in individual wages and labor supply, emphasizing the key differences between major population groups.

Aggregate labor market trends

We start with an overview of aggregate trends in wages and employment. Several studies observed that during the downturn period in Russia, the decline in employment and hours of work was small while the wage decline was large relative to the output decline, in contrast to Central and Eastern European transition economies (Boeri and Terrell 2002, World Bank 2002). We find that the post-1998 economic growth was also accompanied by significant wage adjustments and relatively small changes in employment and working hours.

Hourly real wage level experienced dramatic movements, down 48 percent, or 10 percent per year, during the downturn and up 87 percent, or 9 percent per year, during the recovery (Figure 3A). Panel A of Figure 3 shows actual hourly wage, defined as the ratio of actual labor earnings received last month from all regular jobs to actual hours worked, and compares it to contractual hourly wage (available 1998-2005), which is the ratio of average monthly labor earnings in the last 12 months to usual hours of work per month. The actual wage is higher than contractual wage, partly because actual hours are lower. Male wages appear to be more responsive to output fluctuations: their actual hourly wages declined faster in downturn, but they also grew more rapidly in recovery.

In contrast to wages, hours of work do not vary considerably over time (Figure 3B). Even in the downturn, an average employed person worked more than 40 hours per week. The response of hours to the 1998 financial crisis was minimal. Usual hours of work are relatively high (48 hours in all jobs for males), and they are bigger than actual hours because of temporary absence from work. Females typically work 5-6 hours less per week than males. The share of full-time workers does not change much in response to output fluctuations: it increases slightly over time for both genders, with a somewhat larger overall rise for females during 1994-2005 (Figure 3C).

Employment-to-population ratio in Russia is high by international standards. However, it declined significantly for males from 94-96 percent in 1985-1990 to 86 percent in 1994, and then down to 79 percent in 1998 (RLMS 2000, retrospective questions). In the growth period, the ratio did not reach its pre-crisis levels and stayed relatively constant at 83-84 percent for 25-59 age group (Figure 3D). On average, the employment rate for females is 8 percentage points lower than that for males, which is a smaller gender gap compared to 14 percentage points in the U.S. for the same age group (U.S. Bureau of Labor Statistics 2006). Figure 3D also shows that the official employment rate is lower than that in RLMS in the 1990s, but the difference between the two data sources vanishes in later years.

Earnings and wage inequality

The onset of the transition to a market economy was accompanied by rapidly growing wage inequality (Commander, Tolstopyatenko and Yemtsov 1999). We estimate that the Gini coefficient for earnings increased from 0.28 in 1985 and 0.32 in 1990 to 0.48 in 1995 (RLMS 2000, retrospective questions).[9] The 90/50 ratio climbed from 2.2 in 1990 to 3 in 1995, while the

[9] This dynamics of the Gini coefficient is consistent with other studies. For example, Flemming and Micklewright (2000) report an increase in the Gini coefficient for per capita income from 0.27 in 1989 to 0.41 in 1994 based on the Household Budget Survey. They note, however, that inequality could have been larger in the Soviet period after accounting for significant in-kind subsidies (e.g., free housing).

50/10 ratio rocketed from 2 to 4 in just five years. The latter trend suggests that low-paid workers suffered the most during the first years of market reforms.

During subsequent years, however, wage inequality ceased to grow, as can be seen in Figure 4. This figure compares four different measures of inequality for individual earnings and hourly wages, both actual and contractual, in 1994-2005.[10] Actual earnings received last month are much more variable than contractual earnings. Part of the reason is delayed wage payments, which were a major phenomenon during 1994-1998.[11] Wage arrears tend to exaggerate earnings inequality. For example, some workers in the sample may have received several months of back pay and others received no pay in the reference month, generating income dispersion that is higher than dispersion in annual incomes. At the peak of wage arrears in late 1998, 62 percent of Russian workers reported overdue wages averaging 4.8 monthly salaries per affected worker (Earle and Sabirianova Peter, forthcoming). Consequently, the difference in actual and contractual earnings inequality was the largest in 1996-1998. Wage arrears subsided in later years, although they did not disappear entirely: about 12 percent of all employees reported delays in wage payments in 2005. Because of this, and also due to seasonal and irregular employment, actual earnings still show higher inequality than contractual earnings in later years. In Figure 4, measures of inequality for hourly wages and earnings are close because of low dispersion of working hours.

According to most measures in Figure 4, inequality in wages and individual earnings has been declining over the sample period. The Gini coefficient for contractual earnings declined from 0.48 in 1995 to 0.41 in 2005 and the variance of logs decreased by 0.17. The decline in earnings inequality is more pronounced in the bottom half of earnings distribution: while the

[10] The observations on contractual earnings are available starting in 1998. For 1994-1996, we construct contractual earnings from the data on actual earnings and answers to questions about accumulated overdue wage amount and number of months of overdue pay, according to the method proposed by Earle and Sabirianova (2002). See Appendix 1 for details.

[11] Other reasons for excessive volatility of actual earnings in 1994-1998 include widespread temporary layoffs in the form of unpaid involuntary leaves and forced in-kind payments in lieu of wages owed. The use of involuntary leave peaked in 1996, when 15.8 percent of employees had average leave duration of about eight weeks. In-kind substitutes for money wages peaked in 1998, with 15.4 percent of workers affected (World Bank 2002). Adding these forced substitutes to actual earnings extends the bottom of the distribution of positive last-month income receipts and thus increases the overall dispersion.

90/50 ratio hardly changed over the sample period, the 50/10 ratio fell sizably from 4 to 2.5. Several factors may have contributed to the decline in earnings inequality: oil-driven growth that created labor demand in low-skill industries such as mining and construction, enhanced competition in the labor market (e.g., the number of employers increased tremendously), improved compensation in the public sector, etc.; each of these factors deserve a separate study.

Although the inequality indices remained higher than their pre-transition levels, the overall inequality decline is quite remarkable, and the reasons for it merit further research in the future. This trend is consistent with international macro data showing a negative contemporaneous correlation between income inequality and economic growth for less developed countries (Barro 2000).

Many Russians may be surprised to find that inequality has declined given the emergence of the conspicuous wealthy elite and a popular belief in the rising gap between rich and poor. We note, however, that adding the super-rich to the RLMS data will not affect the Kuznets ratios in Figure 4. There still might be a valid concern that upwardly mobile high earners may have left the addresses surveyed by the RLMS interviewers, and that those who stayed are self-selected low earners. Some of the issues with panel attrition are addressed within the survey itself by adding new dwellings to the sample and adjusting the sample weights.[12] The fact that our sample is unlikely to be skewed towards poor is also supported by the RLMS aggregate expenditure levels that are close to NIPA and exceed the HBS levels reported in Section 2.

Wage premia

The analysis of between-group wage inequality reveals several interesting results. They are reported in Figure 5 that shows aggregate trends in wage premium associated with education, gender, and experience. The male education (college/non-college) premium is substantial (about 50 percent on average), although it is smaller than the current education premium in the U.S.

[12] To assess the importance of non-random exit from the survey on the measures of inequality, we re-weighted observations by giving a larger weight to observations with a higher probability of exit. The adjusted weight is calculated as *L.weight* × 1/(1- *Pexit*), where *L.weight* is the sample weight from the previous round and *Pexit* is the probability of exit from the survey estimated from a flexible probit regression that includes a wide range of controls for individual characteristics. We found that adjustment for non-random exit barely changes the magnitude and the trend slope of earnings inequality.

(e.g., Autor *et al* 2008, Eckstein and Nagypal 2004). The education premium has been rising since 1995 but dropped after 2002.

The gender premium in monthly earnings is large (up to 69 percent in 2000), even though it declined to 51 percent in recent years. The gender differences in hourly wages are smaller (35-47 percent) due to fewer hours of work by females. The level is comparable to the U.S. gender premium in the 1970s (e.g., Blau and Kahn 2000).

It is interesting that the male experience premium is negative, and that it is below the female experience premium (Figure 5C). The age-earnings profile reaches its peak at age 33 for males (44 for females), whereas male earnings growth in the U.S. continues until much later ages (e.g., Heckman *et al* 2008). This unusual earnings profile may be partly attributed to the obsolescence of skills of Soviet-era workers.[13] However, if obsolete skills were the sole driving force of the negative experience premium, one would expect the experience premium to be low at first and to rise gradually over time as the old-era workers move out of the labor force. In fact, the male experience premium stays negative and roughly constant throughout the sample period. Another explanation for this result is that dramatic economic changes during both contraction and growth periods generated a wage premium for younger workers because they are more mobile and more adaptive. We also think that deteriorating health, particularly for males, could be a contributory factor to the negative experience premium. The life expectancy of Russian males has dropped by 6.6 years, from 64.2 to 57.6, in just five years prior to 1994 (Brainerd and Cutler 2005). To the extent that this signals deteriorating health of males in their 50s, the "physical decay" of human capital could drag down the experience premium.

The residual inequality trends down over time, which is expected since the overall inequality is declining while the various wage premia for observable characteristics stay roughly constant (Figure 5D). By way of comparison, the residual wage inequality has an upward trend in the U.S. (e.g., Autor 2008, Lemieux 2006).

[13] Consistent with this, Guriev and Zhuravskaya (2008) find evidence of a big shift in life satisfaction by cohort: individuals who finished their education just before the transition report much lower life satisfaction than similar individuals who finished their education just after. This jump in life satisfaction could, perhaps, reflect brighter lifetime earnings prospects of workers educated under the new regime.

Gender differences in labor market outcomes

Figure 6 presents gender comparisons of inequality in hourly wages and hours worked. Wage inequality is higher among males than females, which is found in the U.S. data as well (e.g., Eckstein and Nagypal 2004). Measures of wage inequality for both genders trend down over time, although the decline in inequality is more pronounced for males (this is again consistent with a higher responsiveness of male wages to output fluctuations). Consequently, the differences in wage inequality between genders become less noticeable by the end of the sample period (Figure 6A). Contractual wages show less dispersion than actual wages for both genders.

Hours worked are considerably less variable than wages (note that panels A and B have different scale), but females have slightly more variable hours, perhaps due to higher prevalence of part-time work. Dispersion of hours falls during 1994-1996 and stays stable afterwards.

The bottom two panels of Figure 6 show the correlations between hours and wages for males and females. These correlations are negative for both genders, which is probably due to a downward bias induced by a measurement error in hours, known as "division bias" (Borjas 1980). There is no clear time trend in the correlation between wages and hours for either gender.

Overall, the observed group differences in labor market outcomes behave in expected ways, with the exception of the negative male experience premium. We now turn to the analysis of inequality across households.

4. Inequality in Household Income and Consumption

This section analyzes the aggregate trends in income and consumption inequality at the household level. We first examine inequality in household labor earnings and then show the contributions to inequality from financial income, private transfers, government transfers, and home production. We also compare income inequality to consumption inequality and discuss possible reasons for the observed differences.

Inequality in household labor earnings

The RLMS data have several sources of information on household labor earnings. Our preferred measure, *yL*, is aggregated from individual responses on after-tax contractual labor earnings (see Appendix 1 for details). We note that Russian households are rather large and often include multiple generations of adults and extended family. The average number of adult members (14+) is 2.6, and it is not rare for a household to have more than two earners (see Appendix 3 for the sample composition of households). In this case one needs to be particularly careful when aggregating individual responses to the household level and should adjust for non-response.[14] However, since the RLMS response rate within the household is fairly high (about 97%), this adjustment does not affect the mean and the variance of labor earnings (e.g., compare *yLc* and *yL* in Figure 7A). Figure 7A shows that over time, the dispersion (var-log) in total contractual earnings across households is trending downward, similar to earnings dispersion across individuals.

Another measure plotted in Figure 7A is the variance of the logarithm (var-log) of household actual labor earnings received last month (*yLa*). These earnings are reported on behalf of all household members by the reference person. While contractual earnings are monetary, actual earnings also contain non-monetary compensation (including forced in-kind substitutes for cash payments) that may introduce additional variability to household earnings. The dispersion in actual labor earnings has been declining after 1998, but its magnitude is considerably higher in comparison to contractual earnings, especially in the second half of the 1990s. As Figure 7B exhibits, the 1996-1998 period have the largest share of working households affected by wage arrears (67-68 percent) and forced in-kind substitutions of payments (17-20 percent), which are the two most likely contributors to high earnings volatility (see also footnote 9). By 2005, the difference in dispersion between actual and contractual earnings reduces significantly, but it does not disappear entirely, possibly due to a measurement bias of one-person reporting, irregular employment, and residual wage arrears.

[14] We impute labor earnings for non-respondents using their demographic characteristics known from the roster files and location.

The variance of the log of labor earnings can be decomposed into parts accounted for by observable components based on the following regression:

$$\ln(yL_{ht}) = \beta_{0t} + \beta_{1t}D_{ht}^{H} + \beta_{2t}D_{ht}^{L} + \beta_{3t}D_{ht}^{E} + f_t(a_{ht}) + u_{ht}\,, \tag{1}$$

where yL_{ht} is contractual labor earnings of household h in year t, β_{0t} is year-specific intercept, D_{ht}^{H} is a set of dummies for household composition (e.g., categories for size, number of children, and number of seniors), D_{ht}^{L} is a vector of location characteristics such as an urban dummy, a dummy for Moscow and St. Petersburg, and 7 dummies for federal districts, D_{ht}^{E} denotes a set of dummies for educational attainment of the head of household, $f_t(a_{ht})$ is a quartic polynomial in age of household head, and u_{ht} is the error term (see Appendix 1 for details on how these components are constructed). The equation is estimated separately for each year. The observables explain a significant portion of inequality; however, the residual inequality remains large (46-62 percent, as shown in Figure 7C). Figure 7D plots the contributions of observable components to the overall dispersion of household labor earnings. Location and household composition factors contribute the most to the observed inequality; education contributes some but age contributes close to zero. Because of its importance for inequality in Russia, we will consider the effect of location on inequality in more detail in Section 6.

To account for household size, we compute the equivalized household labor earnings using the OECD equivalence scale.[15] The dispersion for log equivalized earnings is almost the same as raw dispersion, suggesting that equivalized earnings are negatively correlated with household size (Figure 7C). Figure 8 presents several alternative measures of inequality in household labor earnings per adult equivalent. Similar to the var-log measure discussed earlier, the Gini coefficient and both Kuznets ratios exhibit a downward trend in the recovery period and show rapid convergence in inequality between the actual and contractual measures of labor earnings after 1998.[16]

---

[15] The OECD equivalence scale assigns a value of 1.0 to the head of the household, a value of 0.7 to each additional adult (17+), and a value of 0.5 to each child.

[16] It is interesting that inequality levels are similar for individual labor earnings and equivalized household labor earnings (compare Figures 4 and 8). This similarity, however, is not due to a strong correlation of individual labor earnings within households (<0.3 in our data) but because the denominator is the number of adult equivalents (which

From wages to disposable income

Using the variance of the log, we analyze how income inequality changes as we add different components of household income. Figure 9A shows that the magnitude of dispersion and its trend hardly change as we move from hourly to monthly contractual earnings of household head and then add earnings of other household members (*yL*). The var-log of all three measures of labor earnings is trending downward.

Taking labor earnings as the base, we add income from other sources one at a time and report the corresponding inequality trend in panels B and C of Figure 9. For comparability purposes, all trends are estimated on a consistent sample of working households with non-zero contractual earnings. Net private transfers further increase the dispersion of earnings throughout the whole period (the *yL+* line in panel B is above the *yL* line), conceivably, because they are made irregularly.[17] Income derived from financial assets is negligible for most households, which is why financial income has virtually no effect on inequality. Government transfers, on the other hand, play a significant role in reducing income inequality, especially after 1998 (see *yD* line in panel B). The spike in income inequality in 1996 could be explained by unusually high pension arrears and unemployment benefit arrears in that year. Having income from subsidiary farming at home (which includes both own consumption valued at market prices and sales of home grown food) also have a large equalizing effect on earnings distribution, as evidenced in panel C.[18]

The dispersion of disposable income of working families with one or more wage earners exhibits a downward trend since 1996. However, adding non-working families (about 11% of the sample) not only shifts the overall income inequality up but also alters the time trend (see Figure 9D). This is because non-working families whose income consists of small private or

---

is negatively correlated with the average earnings of household members) rather than the number of working household members.

[17] The correlation between net private transfers and household labor earnings is -0.14, suggesting that that more affluent households are likely to support other households, while not-so-affluent families are likely to receive support from others. However, unlike public transfers, sporadic lump-sum private contributions may cause sizeable movements in the resources available to households, which raise our measures of income inequality. In our view, private transfers would probably decrease inequality if they were measured on annual basis.

[18] A related study by Gottschalk and Mayer (2002) shows that income adjusted for the value of home production is more equally distributed than unadjusted income in the U.S.

public transfers are more likely to fall into the bottom end of income distribution. Over time, more non-working households report small positive income either because they started receiving public transfers or because payment of transfers became more regular. This may explain why income inequality for the pooled sample of working and non-working families does not decrease over time.

As previously discussed in Section 2, income is likely to be under-reported. To the extent that income under-reporting varies by income level, under-reporting can introduce a bias in measures of cross-sectional income inequality. For example, if higher income households report a smaller fraction of their income than the average household, cross-sectional measures of income inequality will be biased downwards. There are also reasons to believe that income under-reporting is time-varying, and it declined following the Russia's 2001 flat tax reform (see Gorodnichenko *et al* 2008 for evidence). If this is the case, then the attenuation of income reporting bias towards the end of our sample period makes the true fall in income inequality even larger than that in Figures 7-9.

From disposable income to consumption

Our benchmark measure of consumption, non-durable expenditures, is constructed from data on food expenditures in the reference week and expenditures on non-food items in the last 30 days (see the overview in Section 2 and Appendix 1 for more details). Figure 10 shows that the dispersion of non-durable consumption increases significantly during the downturn and falls rapidly during the economic recovery.[19] Other consumption variables follow this trend very closely, although their variance may have different magnitude. In particular, adding durable expenditures (*cD*) increases consumption variance while adjusting for services from owned housing reduces it (*cD+*). The equalizing effect of housing on consumption distribution is predictable since many households, especially older ones who are also poorer, inherited their housing from the Soviet era.

[19] Consistent with this, Artamonova *et al* (2007) find that top to bottom expenditure quintile ratio in the RLMS has been falling during 1998-2005, but rose in 2006.

Figure 10 also presents decomposition of consumption variance using the same covariates as in equation (1). The consumption decomposition results are broadly similar to those for earnings (see Figure 7). The dispersion of equivalized consumption is slightly lower than the dispersion of raw consumption. The residual consumption inequality is large and follows the same time pattern as the raw measure of consumption inequality (Figure 10C). As was the case with income decompositions, the largest observable contributors to consumption inequality are household composition and location. Education of household head explains some of the consumption inequality, but age explains almost none (Figure 10D). By contrast, in the U.S. inequality across households typically grows with age. The lack of correlation between measures of inequality and age in Russia is also reflected in the flat life-cycle inequality profiles (see Appendix 4).

Next, we compare consumption inequality to income inequality. Each of the four panels of Figure 11 depicts a different inequality measure: variance of the logs, 90-50 percentile ratio, 50-10 percentile ratio, and the Gini coefficient. Note that the measures of income inequality are for *all* households, including non-working. This makes income inequality measures different from those for earnings (see also Figure 9D). We observe that actual disposable income received last month (*yDa*) is highly volatile for the reasons discussed earlier, i.e., wage arrears, non-monetary compensation, temporary layoffs, etc. As the economy stabilizes, inequality of *yDa* declines considerably and approaches the inequality of disposable income based on contractual labor earnings (*yD*).

We also see that the measures of consumption inequality are much closer to the inequality measures of *yD* than those of *yDa*. In some periods consumption inequality exceeds income inequality (compare lines *c* and *yD*). This could be due to a downward bias in measures of income inequality induced by underreporting among high earners (recall the discussion above). In particular, this can explain why the 90/50 ratio is higher for consumption than for income throughout the entire sample period.

In addition, consumption inequality may be subject to its own biases. For example, consumption inequality was markedly higher than income inequality during 1996-1998 (compare lines *c* and *yD* in Figure 11). High consumption inequality in this particular period may be driven by the tendency to store food as a means of short-term consumption smoothing. Specifically, wage and government transfer arrears made household monthly income highly variable (e.g., see the *yLa* line in Figure 7). In perfect financial markets, these income variations would be smoothed by changing the stock of household financial assets. However, most households in our sample do not hold significant financial assets, perhaps due to undeveloped financial markets or the low real rate of return. We think that short-term consumption smoothing may have been done by adjusting food inventories: households that received several months of back pay purchased large quantities of storable food (i.e., flour, sugar, etc.) for future consumption. In this case we can have households that spend little and consume from their food inventories as well as households that spend a lot on food, but do not consume all of it. Thus, the presence of food storage can make consumption inequality artificially high. Since the size of food inventories is unknown, we cannot disentangle food consumption and food expenditures.

There may be another source of upward bias in consumption inequality that has to do with food grown at home. As noted in Section 2, many Russian households grow food on subsidiary plots and thus consume more food than their expenditure numbers suggest. Since food production tends to concentrate among rural and poorer households, expenditure inequality will overstate the true consumption inequality. Adjusting consumption for home-grown food produces a large equalizing effect on consumption distribution for all four measures of inequality, as can be seen in Figure 11, line *cH*. The impact of home-grown food on consumption inequality is particularly large at the lower end of the consumption distribution (compare panel B to panel C).

Thus, wage arrears, income underreporting, food storage, and food production at home may explain some of the differences between trends in income inequality and consumption inequality. Keeping these sources of differences in mind, we will next investigate to what extent changes in income inequality translate into changes in consumption inequality.

5. The Interaction of Income and Consumption Inequality

To understand the dynamics of inequality and the interactions between consumption and income, we need to identify the sources of uncertainty faced by households and to assess households' ability to smooth the shocks.

As a first pass, we exploit the panel aspect of RLMS and decompose the residual variability in consumption and income into permanent and transitory components. Specifically, we use a statistical model $\ln(s_{ht}) = X_{ht}\beta + u_{ht}^{(s)}$, where s_{ht} is the variable of interest, such as income or consumption, and X_{ht} is the same set of controls as in equation (1). We decompose the error term $u_{ht}^{(s)}$ into the sum of a transitory component and a permanent component that follows a random walk process:

$$u_{ht}^{(s)} = \alpha_{ht} + \varepsilon_{ht}, \tag{2}$$

$$\alpha_{ht} = \alpha_{h,t-1} + \eta_{ht},$$

where $\varepsilon_{ht} \sim (0, \sigma_{\varepsilon,t}^2)$ is the transitory shock, $\eta_{ht} \sim (0, \sigma_{\eta,t}^2)$ is the permanent shock, and s_{ht} is a measure of income and consumption. Note that the variances of the transitory and permanent shocks are time-varying. Using covariances of growth in s_{ht} and an equally weighted minimum distance estimator, we find the time series for $\sigma_{\varepsilon,t}^2$ and $\sigma_{\eta,t}^2$ (Figure 12).[20]

Irrespective of what measure of income we use, we observe that the variance of permanent shocks remained relatively stable while the variance of transitory shocks declined considerably.[21] Thus the fall in residual income and consumption inequality is primarily due to moderation of transitory shocks. The variance of permanent and transitory shocks is at least three times larger in Russia than in the U.S. (Meghir and Pistaferri, 2004).

We find a very similar decomposition of shocks to consumption (see Figure 12D). This pattern sharply contrasts with recent trends in consumption and income inequality in the U.S. where dramatically increased income inequality did not translate into large increases in consumption inequality. Such divergence between the two inequality measures in the U.S. has

[20] Due to identification constraints, we impose $\sigma_{\eta,2005}^2 = \sigma_{\eta,2004}^2$, $\sigma_{\varepsilon,2005}^2 = \sigma_{\varepsilon,2004}^2$, and $\sigma_{\varepsilon,1994}^2 = \sigma_{\varepsilon,1995}^2$.

[21] In addition to individual and household labor earnings and disposable income reported in Figure 12, we find that other income measures such as household earnings with income from home production have a similar trend.

been explained by developments in financial markets that allow more risk sharing and consumption smoothing (Krueger and Perri 2006) and by the changes in the persistence of income shocks (Blundell *et al* 2008). Russia witnessed significant advancements in financial markets (especially, consumer credit) towards the end of our sample period, yet we do not observe the divergence between consumption and income variance decompositions.[22] The similarity of variance decomposition for income and consumption is even more remarkable given that Russian households had a variety of consumption smoothing tools such as saving, food storage, home production, variable labor supply, and extended family. The negative correlation between wages and hours and low savings is also consistent with the lack of insurance against income shocks (Heathcote *et al* 2007).

To look at possible changes in consumption smoothing patterns, we examine the response of consumption to permanent and transitory income shocks. We model the sensitivity of consumption to income shocks as in Blundell *et al* (2008):

$$\Delta u_{ht}^{(c)} = \phi_t \eta_{ht} + \psi_t \varepsilon_{ht} + \xi_{ht} - \xi_{h,t-1},$$

where η_{ht} is the permanent income shock, ε_{ht} is the transitory income shock, and $\xi_{ht} \sim (0, \sigma_{\xi,t}^2)$ absorbs measurement errors and non-income shocks to consumption. The coefficients ϕ_t and ψ_t are called "loadings" and measure the responsiveness of consumption to various income shocks. Blundell *et al* (2008) interpret loadings close to one as indicating the lack of insurance against income shocks. In contrast, if loadings are close to zero, then households have enough instruments to insulate consumption from income shocks. Loadings between zero and one can be interpreted as partial insurance.[23]

Assuming that $\Delta u_{ht}^{(yD)}$ follows the process in equation (2), we estimate the consumption response to income shocks, ϕ_t and ψ_t, as well as variances of shocks from the variance-covariance and autocovarаince matrices of consumption and income growth. The coefficients ϕ_t and ψ_t are reported on Figure 13A. We find that the response of consumption to transitory

[22] Consumer credit more than doubled every year between 2002 and 2006 (Goskomstat 2006c).

[23] Blundell *et al* (2008) show that under certain restrictions the permanent income hypothesis implies ϕ=1 and ψ=0. That is, consumption should change by the same percentage as the change in the permanent income (ϕ=1), and it should not respond at all to transitory shocks (ψ=0).

shocks to disposable income fell from approximately 0.15 in 1995 to 0.03 in 2005, which is consistent with household being increasingly able to smooth transitory income shocks. Likewise, the coefficient on the permanent income shock decreased from 0.9 in 1995 to 0.7 in 2005. Interestingly, the 2005 estimates of ϕ_t and ψ_t are close to those reported in Blundell *et al* (2008) for the U.S. households. The weaker response of consumption to permanent shocks is, in theory, consistent with better self-insurance.[24]

Using our estimates, we can compute the variance of consumption growth that is due to income shocks,

$$\sigma_{cy,t}^2 = \phi_t^2 \sigma_{\eta,t}^2 + \psi_t^2 \sigma_{\varepsilon,t}^2 .$$

This variance stayed relatively constant and modest, with $\sigma_{cy,t}^2 = 0.04$. In contrast, the variance in the growth of non-durable consumption due to non-income shocks ξ_{ht} was substantially larger: it started at 0.25 in 1995 and declined to 0.14 in 2005. Hence income shocks can only explain between 14 and 22 percent of variance in non-durable consumption growth.

Given that incomplete insurance against income shocks appears to be a salient feature of Russian reality, one may be interested in assessing the benefit of providing such insurance to households. One can use the simple framework of Lucas (1987) to obtain a rough estimate of the welfare gains from eliminating fluctuations in consumption. Assuming relative risk-aversion parameter $\gamma = 4$, which is consistent with Barsky *et al* (1998), we find that households would be willing to sacrifice up to $s \approx \frac{1}{2}\gamma\sigma_{cy,t}^2 = \frac{1}{2} \times 4 \times 0.04 = 0.08$ share of their consumption to eliminate income shocks that are not smoothed out. Obviously, with lower risk aversion, one can reach a smaller cost of consumption fluctuations. However, the cost of consumption fluctuations is likely to remain high given that consumption fluctuates significantly in response to non-income shocks.

Consistent with the permanent-transitory shock decompositions above, we find that the variance of the growth of income and consumption declined over 1995-2005 (Figure 13B).

[24] This result has an important caveat: if households smooth consumption with food storage, this would induce strong response of expenditure to transitory income shocks, but will not necessary imply non-smooth consumption. Because of this, it is hard to tell whether Russian households did in fact achieve better consumption smoothing by the end of the sample period.

Interestingly, the variation in observable characteristics of households contributed little to explaining the decreasing variance of consumption and income growth (Figure 13C).[25] In a similar vein, we find that the probability of staying in the same income quintile increased by 5 to 10 percentage points over 1995-2005 (Figure 13D).[26] Hence, the ranking of households in terms of income has been stabilizing in recent years.

Overall, our findings suggest that mobility in the early stages of transition was high and there was significant randomness in who lost or won during the 1990s. Despite recent improvements, households have had limited ability to smooth income shocks with financial assets, savings or other insurance instruments and the benefit from providing access to such insurance is probably substantial.

6. A Closer Look at Inequality Trends: the Role of Location and Prices

In the context of the Russian economy, two factors deserve special consideration as they can help with understanding of the observed inequality trends. This section takes a closer look at the role of geography and price dispersion.

Location effect

Income and consumption decompositions in Section 4 point out that location is the most important explanatory variable for the dispersion of equalized earnings and consumption (see Figures 7 and 10). This is not surprising, given how large and diverse Russia is. For example, monetary income per capita in the richest Russian region is 10.6 times larger than per-capita income in the poorest region in 2005 (Goskomstat 2007b). A similar maximum-to-minimum ratio across states in the U.S. is only 1.8 (U.S. Census Bureau 2007).

The substantial dispersion of the regional component of inequality may be associated with the large geographic variation in the cost of living. The 2005 ratio in the cost of fixed

[25] The result could be due to our choice of observables that has low variance over time. Other observable characteristics pertaining to sectoral shifts and changes in jobs and occupations may be better predictors of income and consumption growth but we will leave it for future research.
[26] This probability is still less than its counterpart in the U.S., where it is in the 0.5-0.6 range (Moffitt and Gottschalk 1995).

consumer goods between the most expensive region and the least expensive region was 2.7 (Goskomstat 2006a). With such inter-regional diversity, using a common national CPI may overstate the extent of inequality in both income and consumption. Indeed, using regional CPI and accounting for the regional differences in the cost of living move the magnitude of inequality down, but this adjustment does not affect the time trend (see Figure 14A).

The regional dispersion of expenditure may also be affected by uneven distribution of amounts of food grown at home between urban and rural households. While big city residents purchase more than 95 percent of their food at the store, residents of small towns and villages purchase about 80 percent at the store (less in early years) and grow the rest on their subsidiary plots. Consequently, rural households are likely to have a larger discrepancy between expenditure inequality and consumption inequality. Panels B and C of Figure 14 implicitly confirm this argument. The panels depict the variance of the log of non-durable expenditures for the two groups and the pooled sample using regional deflators, with and without food grown at home. Expenditure inequality is apparently much higher among the rural population (Panel B). By contrast, inequality in *consumption* that includes food purchased and grown at home is much more similar across urban and rural households (panel C).

While time trends in expenditure inequality for the two groups are similar, trends in consumption inequality diverge during economic recovery. In particular, consumption inequality among rural households shows no downward trend (panel C). This difference in trends is consistent with transition of rural households from subsidiary farming to professional farming, which could have made the amount of food grown for own consumption more unequally distributed.

It is interesting to see if economic consequences of downturn and recovery differed between urban and rural populations. One would expect rural households to fall behind during the transition due to the lack of access to large and diverse labor markets that big cities offer and also because of possible migration of the ablest workers to cities.

Surprisingly, our data do not point to much divergence in the mean levels of income and consumption of the two groups until 2002. Figure 14D shows that the relative levels of disposable income (*yD*), expenditures (*c, cD*) and consumption (*cH*) stay fairly constant during 1994-2001. In other words, there is no evidence that rural areas experienced a relatively more severe economic downturn. On the other hand, the relative consumption level of urban household was at its all-time high in 2002, 2004 and 2005, suggesting that rural households did lag behind as the economic recovery progressed. Particularly, the growth of durable consumption was stronger among urban households (*c* and *cD* lines in Figure 14D diverge after 2000).

The role of food grown at home in equalizing consumption is strikingly apparent in Figure 14D. Urban households, who spend 45 percent more than rural households, enjoy only 29 percent higher consumption, on average (compare *c* and *cH* lines).[27]

Comparisons of group income and consumption differences reveal important facts. On average, urban households report roughly 71 percent more disposable income than rural households, but their total expenditure is just 45 percent higher. Because saving rates of most households are fairly low, this leads us to suspect that income under-reporting is more severe among the rural households.

Price dispersion effect

Data on food prices available in RLMS allow us to investigate the effect of prices on the mean and the variance of real expenditures. It is important to examine price data for two reasons: (1) there is a concern that the growth rates of real output and expenditures may be mismeasured due to a bias in the official CPI (Gibson *et al* 2004), (2) expenditure inequality can arise as a result of either price dispersion or quantity dispersion, and only the latter captures inequality in actual consumption.

Using the data on food prices, we find that the official CPI substantially overstates inflation during 1994-1998. We also find that the effect of price dispersion of inequality is

[27] Market value of home production probably overstates its net contribution to household welfare because of the cost of capital goods and materials and decreased leisure. Selling food is also likely to involve high transaction costs, making net income from home production lower than its value at market prices.

relatively small: inequality measures based on expenditures and those based on quantities purchased are close to each other.

Figure 1C shows that food expenditure fell faster than income during the 1994-1998 downturn (see Figure 1B and also the discussion in Mroz *et al* 2005), but never returned to its 1994 level during the economic recovery. Since food expenditures declined rapidly during the early, high inflation, period, one may suspect that the decline is not genuine and may be driven by a bias in the national CPI that we use to deflate food expenditures.

To check this, we construct the food CPI from RLMS prices and compare it to the NIPA deflator for food. Let p_{kt} denote the sample average unit price of food category k in year t, and q_{kt} denote the average physical quantity of food item k purchased in year t. Let $\overline{p}_k$ and $\overline{q}_k$ be the sample average price and the quantity purchased in the base year. Define a fixed-basket food CPI as

$$cpiF_{RLMS,t} = \sum_k p_{kt}\overline{q}_k \Big/ \sum_k \overline{p}_k\overline{q}_k$$

Figure 15A depicts the year-on-year growth rates of $cpiF_{RLMS,t}$ (with base year 2002) and the NIPA CPI deflator.[28] Unfortunately, the two deflators are not directly comparable, because we could not replicate the NIPA procedure without knowing the NIPA choice of base year and when it was changed. Predictably, two deflators in Figure 15A disagree most during the years with high inflation: the inflation rate derived from RLMS data on food prices is lower than the official CPI inflation in 1995 and 1998 years. As a result, between December 1994 and December 1998 food prices in RLMS grew by a factor of 4.5, compared to a factor of 5.6 according to NIPA. If the RLMS food deflator was used in place of the official CPI to compute real consumption, the 1998 value of aggregate consumption (*cD*) in Figure 1 would have been almost 25 percent higher.

The CPI deflator can also be subject to substitution bias if the composition of food consumption changes. To check whether this is the case, we compare two alternative measures of real food expenditures: nominal expenditures deflated by the fixed-basket food CPI, $cF_{RLMS,t}$

[28] To make it comparable, we calculate the NIPA inflation rate as the average monthly inflation rate weighted by the share of respondents interviewed in a given month.

$$cF_{RLMS,t} = \sum_k p_{kt} q_{kt} \Big/ cpiF_{RLMS,t}$$

and the food quantity index, $qF_{RLMS,t}$

$$qF_{RLMS,t} = \sum_k \overline{p}_k q_{kt}$$

that weights current year quantities at base year prices. By construction, the ratio of the two expenditure measures equals the ratio of the CPI derived from the current year basket to the CPI derived from the base year basket. If $cF_{RLMS,t}$ and $qF_{RLMS,t}$ are substantially different, this is an indication of a time-varying consumption basket.

Figure 15B compares $cF_{RLMS,t}$ and $qF_{RLMS,t}$ and find that they are very similar to each other, suggesting that the food consumption basket was essentially fixed throughout the sample period. However, the real food consumption cF (computed with the official deflator) is substantially higher than both $cF_{RLMS,t}$ and $qF_{RLMS,t}$ during 1994-1998, indicating inflation overstatement by the official CPI.

Despite their apparent differences, all measures of real food expenditure show a decline over the sample period. The share of food in aggregate expenditures, a measure not dependent on the definition of deflator, also steadily declined over the sample period (Figure 15C). At the same time, the ratio of NIPA food deflator to NIPA non-food goods deflator (that excludes services), a proxy for the relative price of food, has not changed much (Figure 15D).

From a viewpoint of static utility maximization, a change in demand may be driven either by a change in income or a change in relative price. We do not think that income change was driving the food demand, because real income was roughly the same in 1994 as in 2002, but food expenditures were much lower in 2002 than in 1994. Superficially, at least, there is no change in the relative price of food either, so the apparent fall in the share of food expenditures is puzzling.

We think that one explanation for the falling share of food may be the lack of quality adjustment in the NIPA non-food CPI. The Soviet-era consumer goods were notorious for their low quality. If the relative quality of non-food goods rose over the sample period, this may have caused the quality-adjusted goods price to fall and the expenditure share to shift away from food towards non-food goods.

To summarize, our examination of food price and quantity data and its comparison with NIPA price indices point to evidence of quality bias, but not substitution bias. Gibson *et al* (2004) use RLMS food expenditure data to indirectly infer the total CPI bias from Engel curves and estimate that 2001 real GDP level may be understated by as much as 30 percent.

It is also important to check if the observed level of inequality in food expenditures arises from the difference in food prices that households face. Russia is a geographically diverse country with large variations in the price level by location. Other peculiar features of the Russian transition, including price liberalization, hyperinflation spikes, regional disintegration, imperfect markets, and elevated uncertainty, may contribute to relatively high price dispersion not only across locations but also within locations. As the economy stabilizes and markets develop, we may expect a decline in the level of price dispersion. On the other hand, the Soviet era products were fairly standardized with low quality variance. Over time, import penetration and domestic competition have brought new products of various qualities, thus increasing price dispersion. The resulting effect of transition on price dispersion is thus ambiguous. Figure 16A shows the trends in dispersion of food prices, overall and within location. Price dispersion was high in 1994 but stayed constant afterwards, suggesting that counter-factors of dispersion cancel each other out.

To control for cross-sectional price dispersion, we measure the inequality in food quantity index, $qF_{RLMS,t}$, that weights the quantity of food purchased at constant, base year prices. Figure 16B compares the variances of $ln(qF_{RLMS,t})$ and $ln(cF)$. Quantity dispersion does not have to be lower than expenditure dispersion, because the prices paid by individual households and the quantity of food that they purchased may be negatively correlated. In fact, Figure 16B shows that quantity dispersion is slightly higher than expenditure dispersion in 1995 and 1996 and slightly lower than expenditure dispersion in other years. Overall, the inequality trends for food expenditures and food quantity are similar.

Expenditure inequality may also be affected by regional differences in the cost of living. To control for this, we deflate food expenditures by *regional* CPI. The *cF-reg* line of Figure 16B shows the resulting measure of expenditure inequality. Predictably, using regional price adjustment has an equalizing effect on consumption distribution. Food expenditures with regional deflators show *less* dispersion than the food quantity index $qF_{RLMS,t}$, which indicates a negative correlation between the region-specific food prices and the quantity purchased.

7. Conclusions and Open Questions

We investigate the levels and the time trends of consumption and income inequality in Russia. The paper makes a number of contributions on issues of inequality measurement. We explain, for example, why consumption that includes home production, avoids underreporting of resources available to households, and is adjusted for regional variation in the cost of living should be a preferred inequality measure for Russian economy. We find that compared to its pre-transition level, inequality first rose and subsequently fell. The rise in inequality appeared to have happened during the price liberalization in the early 1990s while the fall started after 2000. The level of inequality in Russia is now very similar to that in the U.S. (e.g., Krueger and Perri 2006).

We uncover several important facts about inequality in Russia. First, poor households appear to gain from recent economic growth. Second, changes in key observable characteristics of households have a small contribution to the dynamics of consumption and income inequality. The size of permanent and transitory shocks is much larger in Russia than in developed countries. It is fair to say that determination of losers and winners had a significant random component. Third, recent moderation in consumption and income inequality and mobility appears to be driven by the decline in the volatility of transitory shocks. Fourth, consumption and income inequality have similar magnitudes over last 10 years which sharply contrasts with recent dynamics of consumption and income inequality in developed countries. Russian households seem to have limited ability to smooth consumption in face of income and non-income shocks. There are probably large gains from introducing insurance schemes to smooth consumption fluctuations.

Our results also point out some inconsistencies between RLMS and NIPA. In particular, comparisons of consumption levels across data sources suggest that there may be an insufficient adjustment for shadow economic activity in the official statistics. The growth rate of consumption in NIPA has recently become higher than that in RLMS, a phenomenon that was noted in other developing economies (e.g., Deaton 2005). The comparison of CPI levels reveals that NIPA may significantly overstate inflation, and that quality bias is potentially important.

Our analysis highlights several phenomena that merit further research. For example, it is unclear if the falling variance of transitory shocks reflects a time-varying measurement error that gets smaller over time. The finding that income shocks explain a modest part of consumption variance is also puzzling, but the latter does not have to be due to a measurement error. This is because financial markets as well as consumer durables can play a role in consumption smoothing. There is a theoretical possibility that income shocks are mostly absorbed by durable consumption. That is, non-durable consumption may be insulated from shocks, because households react to income shocks by changing the timing and the size of planned durable purchases (e.g., Leahy 2005 and Stacchetti and Stolyarov 2007). The panel structure of RLMS provides a natural data set for investigating the role of durable expenditure as a propagation mechanism for income shocks.

References

Aivazian, Serguei and Stanislav Kolenikov, 2001. "Poverty and Expenditure Differentiation of the Russian Population," *EERC Working Paper Series*, No.01/01.

Artamonova Elena, Irina Denisova, Zlata Dorofeeva, and Marina Kartseva, 2007. "Monitoring Economic Conditions in the Russian Federation: The Russia Longitudinal Monitoring Survey 1992-2004," Report submitted to the U.S. Agency for International Development. Carolina Population Center, University of North Carolina at Chapel Hill, June 2007.

Attanasio, Orazio, Erich Battistin, and Hidehiko Ichimura, 2004. "What Really Happened to Consumption Inequality in the US?" *NBER Working Paper*, No. 10338.

Attanasio, Orazio and Steven Davis, 1996. "Relative Wage Movements and the Distribution of Consumption," *Journal of Political Economy* 104 (6): 1227-1262.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney, 2008. "Trends in U.S. Wage Inequality: Revising the Revisionists," *Review of Economics and Statistics* 90(2): 300-323.

Barro, Robert J., 2000. "Inequality and Growth in a Panel of Countries," *Journal of Economic Growth* 5(1): 5–32.

Barsky, Robert B., F. Thomas Juster, Miles S. Kimball, and Matthew D. Shapiro, 1998. “Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study,” *Quarterly Journal of Economics* 112(2), 537-579.

Blau, Francine D. and Lawrence M. Kahn, 2000. "Gender Differences in Pay," *Journal of Economic Perspectives* 14(4): 75-99.

Blundell, Richard and Ian Preston, 1998. “Consumption Inequality and Income Uncertainty,” *Quarterly Journal of Economics* 113(2): 603-640.

Blundell, Richard, Luigi Pistaferri, and Ian Preston, 2008. “Consumption Inequality and Partial Insurance,” *American Economic Review*, forthcoming.

Boeri, Tito and Katherine Terrell, 2002. “Institutional Determinants of Labor Reallocation in Transition,” *Journal of Economic Perspectives* 16(1): 51-76.

Borjas, George J., 1980. “The Relationship between Wages and Weekly Hours of Work: The Role of Division Bias,” *Journal of Human Resources* 15(3): 409-423.

Brainerd, Elizabeth and David M. Cutler, 2005. “Autopsy on an Empire: Understanding Mortality in Russia and the Former Soviet Union,” *Journal of Economic Perspectives* 19(1): 107-130.

Chen, Shaohua, and Martin Ravallion, 1996. “Data in Transition: Assessing Rural Living Standards in Southern China,” *China Economic Review* 7(1), 23-56.

Commander, Simon, Andrei Tolstopiatenko, and Ruslan Yemtsov, 1999. “Channels of Redistribution: Inequality and Poverty in the Russian Transition,” *Economics of Transition* 7(2): 411–447.

Cutler, David M. and Lawrence F. Katz, 1992. “Rising Inequality? Changes in the Distribution of Income and Consumption in the 1980's,” *American Economic Review Papers and Proceedings* 82(2): 546-551.

Deaton, Agnus, 1997. *The Analysis Of Household Surveys: A Microeconometric Approach to Development Policy*. Baltimore: The John Hopkins University Press.

Deaton, Angus, 2005. “Measuring Poverty in a Growing World (or Measuring Growth in a Poor World),” *Review of Economics and Statistics* 87(1): 1-19.

Earle, John, and Klara Sabirianova, 2002. “How Late to Pay? Understanding Wage Arrears in Russia,” *Journal of Labor Economics* 20(3): 661-707.

Earle, John, and Klara Sabirianova Peter, 2008. “Complementarity and Custom in Wage Contract Violation,” *Review of Economics and Statistics*, forthcoming.

Eckstein, Zvi and Nagypal, Eva, 2004. “The Evolution of U.S. Earnings Inequality: 1961–2002,” *Federal Reserve Bank of Minneapolis Quarterly Review* 28(2): 10–29.

Flemming, John S. and John Micklewright, 2000. “Income Distribution, Economic Systems and Transition,” in Anthony B. Atkinson and Francois Bourguignon, eds., *Handbook of Income Distribution*, Elsevier, pp. 843-918.

Galbraith, James K., Ludmila Krytynskaia, and Qifei Wang, 2004. “The Experience of Rising Inequality in Russia and China during the Transition,” *The European Journal of Comparative Economics*. 1(1): 87-105.

Gibson, John, Steven Stillman, and Trinh Le, 2004. “CPI Bias and Real Living Standards in Russia During the Transition,” *William Davidson Institute Working Paper*, No. 684.

Gorodnichenko, Yuriy, Jorge Martinez-Vazquez, and Klara Sabirianova Peter, 2008. “Myth and Reality of Flat Tax Reform: Micro Estimates of Tax Evasion Response and Welfare Effects in Russia,” *NBER Working Paper*, No. 13719.

Goskomstat, 1999. *Social Situation and Standards of Living of Population in Russia (Sotsial'noe Polozhenie i Uroven' Zhizni Naseleniia Rossii)*. Moscow: The Federal Bureau of State Statistics of the Russian Federation.

Goskomstat, 2006a. *Prices in Russia (Tseny v Rossii)*. Moscow: The Federal Bureau of State Statistics of the Russian Federation, also earlier editions in 1996, 1998, 2000, 2002, and 2004.

Goskomstat, 2006b. *Statistical Methods (Metodologicheskie Polozhenia po Statistike)*, also earlier edition in 1996.

Goskomstat 2006c. Finances in Russia (*Finansy Rossii*), also earlier edition in 2004.

Goskomstat, 2007a. *National Accounts of Russia (Natsional'nye Scheta Rossii)*. Moscow: The Federal Bureau of State Statistics of the Russian Federation, annual series, also earlier editions in 1995-2006.

Goskomstat, 2007b. *Social Situation and Standards of Living of Population in Russia (Sotsial'noe Polozhenie i Uroven' Zhizni Naseleniia Rossii)*. Moscow: The Federal Bureau of State Statistics of the Russian Federation, annual series, also earlier editions in 1996-2006.

Gottschalk, Peter, and Susan Mayer, 2002. “Changes in Home Production and Trends in Economic Inequality,” in D. Cohen, T. Piketty, and G. Saint-Paul, eds., *The Economics of Rising Inequalities*, Oxford University Press.

Guriev, Sergei M. and Andrei A. Rachinsky, 2006. “The Evolution of Personal Wealth in the Former Soviet Union and Central and Eastern Europe,” *UNU-World Institute for Development Economics Research Paper*, No. 2006/10.

Guriev, Sergei and Ekaterina Zhuravskaya, 2008. “(Un)Happiness in Transition,” available at SSRN: http://ssrn.com/abstract=1077709.

Heckman, James J., Lance J. Lochner, and Petra E. Todd, 2008. "Earnings Functions and Rates of Return," *NBER Working Papers*, No. 13780.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni K. Violante, 2007. “Consumption and Labor Supply with Partial Insurance: An Analytical Framework,” *CEPR Discussion Paper*, No. 6280.

Jovanovic, Branko, 2001. “Russian Roller Coaster: Expenditure Inequality and Instability in Russia, 1994-1998,” *Review of Income and Wealth* 47(2), 251-271.

Krueger, Dirk and Fabrizio Perri, 2006. “Does Income Inequality Lead to Consumption Inequality? Evidence and Theory,” *Review of Economic Studies* 73(1): 163-193.

Lemieux, Thomas, 2006. “Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?” *American Economic Review* 96(3): 461-498.

Leahy, John V. 2005. "The Timing of Purchases and Aggregate Fluctuations," *Review of Economic Studies* 72(4): 1127-1151.

Lucas, Robert E., 1987. *Models of Business Cycles*. London: Blackwell.

Meghir, Costas, and Luigi Pistaferri, 2004. “Income Dynamics and Heterogeneity,” *Econometrica* 72(1), 1-32.

Milanovic, Branco, 1999. “Explaining the Increase in Inequality during Transition,” *Economics of Transition* 7(2):299–341.

Moffitt, Robert, and Peter Gottschalk, 1995. Trends in the Variances of Permanent and Transitory Earnings in the U.S. and Their Relation to Earnings Mobility. Manuscript.

Mroz, Thomas, Laura Henderson, and Barry Popkin, 2005. “Monitoring Economic Conditions in the Russian Federation: The Russia Longitudinal Monitoring Survey 1992-2004,” Report submitted to the U.S. Agency for International Development. Carolina Population Center, University of North Carolina at Chapel Hill, April 2005.

Shleifer, Andrei and Daniel Treisman, 2005. “A Normal Country: Russia after Communism,” *Journal of Economic Perspectives* 19(1): 151-174.

Slesnick, Daniel T., 2001. *Consumption and Social Welfare. Living Standards and Their Distribution in the United States*. Cambridge: Cambridge University Press.

Stacchetti, Ennio and Dmitriy Stolyarov, 2007. “Obsolescence of Durables and Optimal Consumption,” *working paper, University of Michigan*, available at http://www-personal.umich.edu/~stolyar/Papers/odoc2006.pdf.

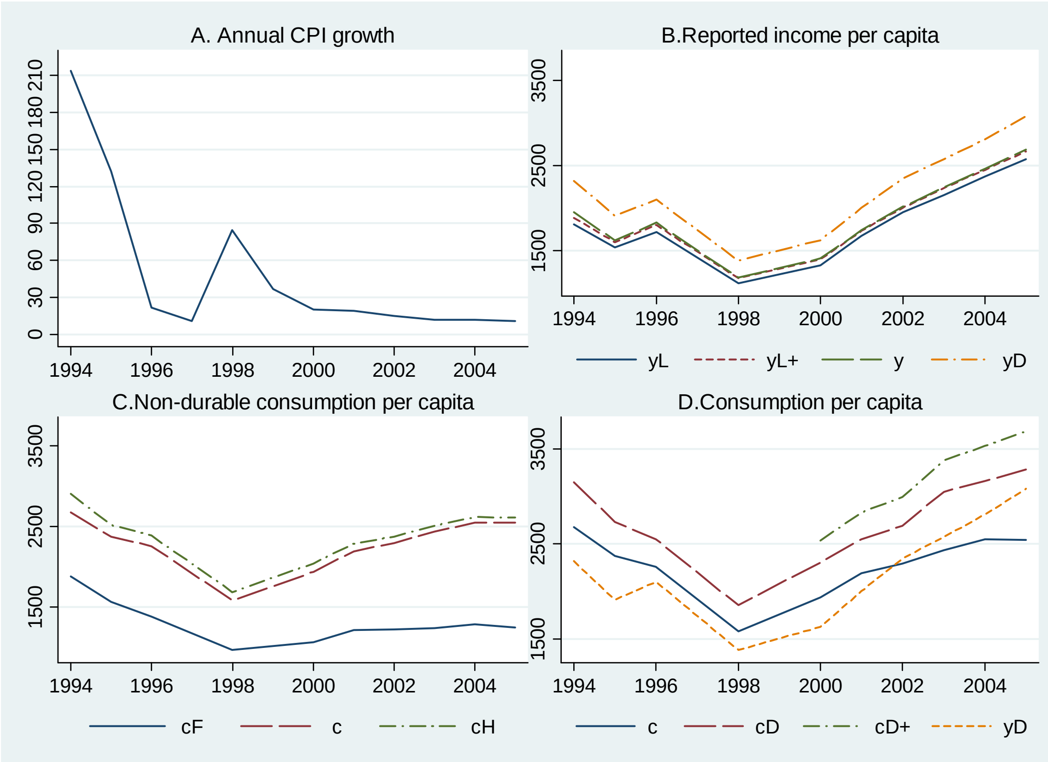
Stillman, Steven, 2001. “The Response of Consumption in Russian Households to Economic Shocks,” *William Davidson Institute Working Paper Series*, No. 412.

U.S. Bureau of Labor Statistics (2006). *Women in the Labor Force: A Databook*. U.S. Department of Labor, Bureau of Labor Statistics, available at http://www.bls.gov/cps/wlf-databook2006.htm

U.S. Census Bureau, 2007. *Statistical Abstract of the United States*. Washington, D.C.

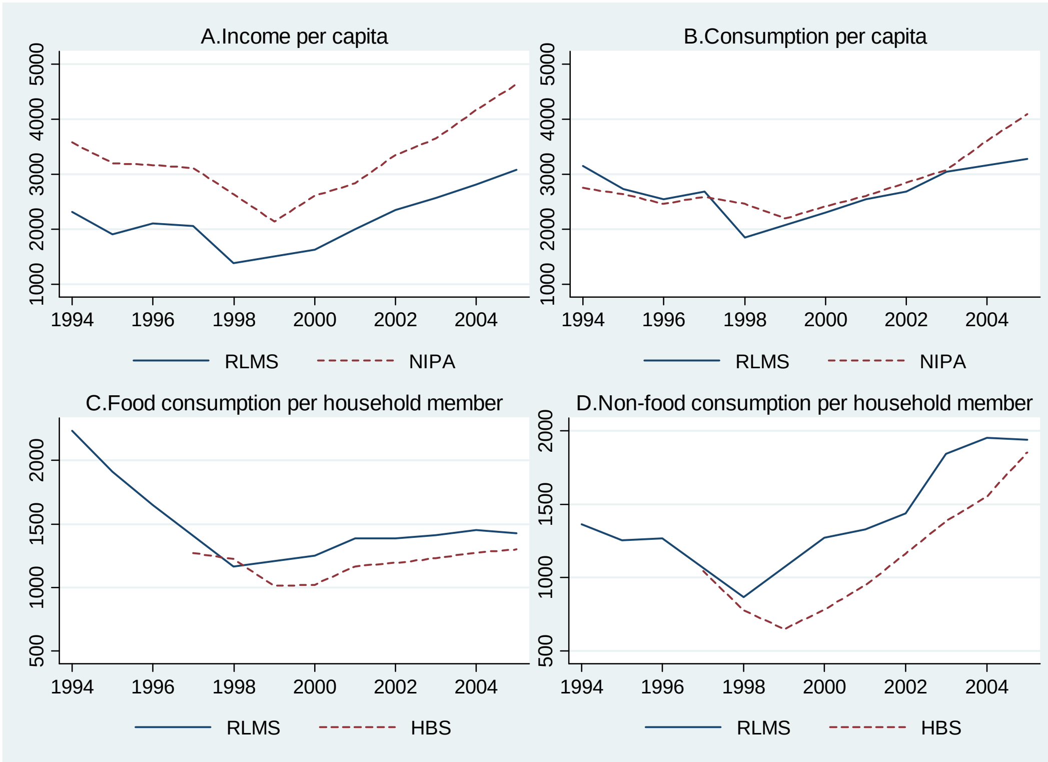
World Bank, 2002. *The Russian Labor Market: Moving from Crisis to Recovery,* Mansoora Rashid, ed., Washington D.C.

Figure 1: Trends in Household Income and Consumption



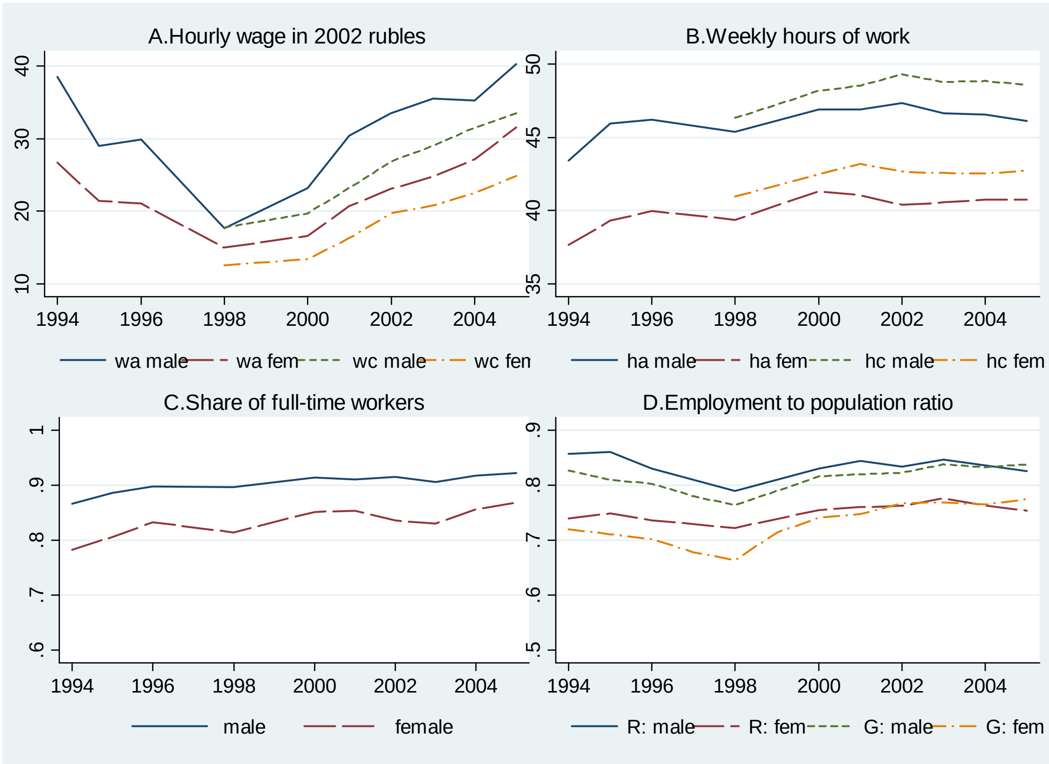
Notes: Panel A shows annual inflation rate using national end-year CPI from official sources. In remaining panels, all measures are in constant December 2002 prices (deflated using national monthly CPI and the date of interview). *yL* = household contractual labor earnings per month; *yL+* = *yL+* net private transfers; *y* = (*yL+*) + financial income; *yD* = disposable household income = *y* + government transfers; *cF* = expenditures on food, beverages, and tobacco last week (multiplied by 30/7); *c* = household non-durable expenditures last month; *cH* = *c* + consumption of home-grown food; *cD* = *c* + expenditures on durables; *cD+* = *cD* + imputed services from housing.

Figure 2: Comparison of RLMS with Official Statistics



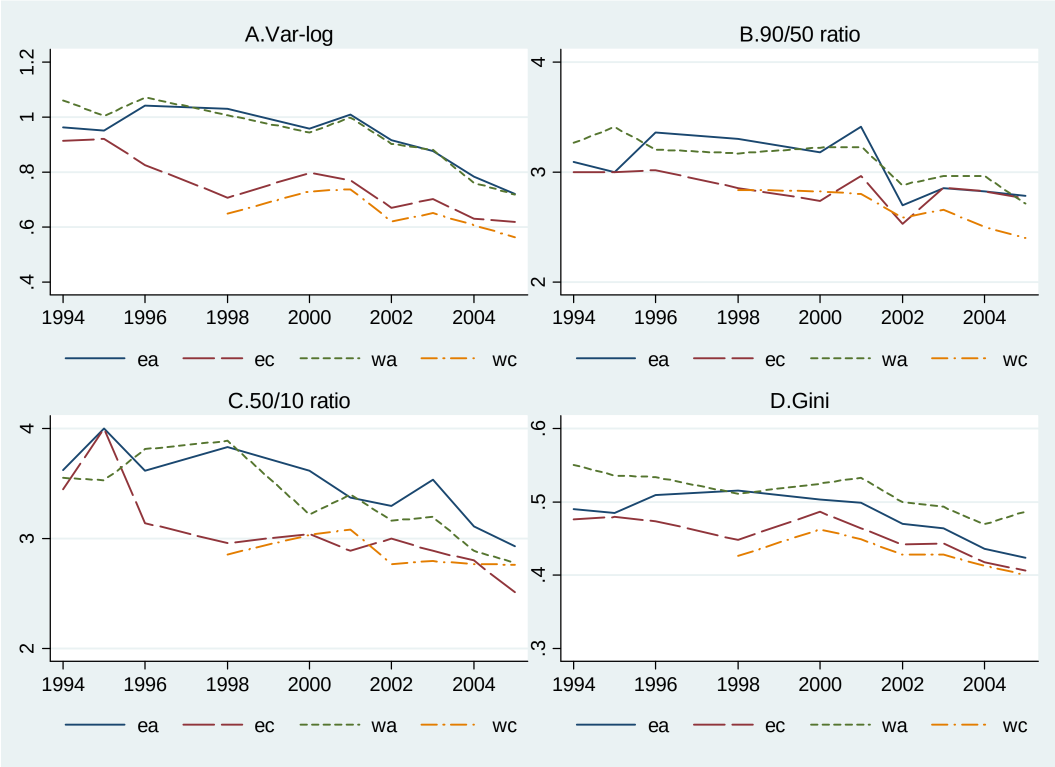
Notes: For comparability purposes, the following RLMS measures are selected: *yD* in panel A, *cD* in panel B, *cF* + consumption of home-grown food in panel C, *cD* - *cF* in panel D. The RLMS sample is unrestricted. All RLMS measures are deflated using monthly CPI and the date of interview. All NIPA and HBS measures are deflated using annual average CPI. RLMS income and consumption for 1997 are imputed using the lagged RLMS value multiplied by the 1997 growth rate from NIPA.

Figure 3: Trends in Labor Supply



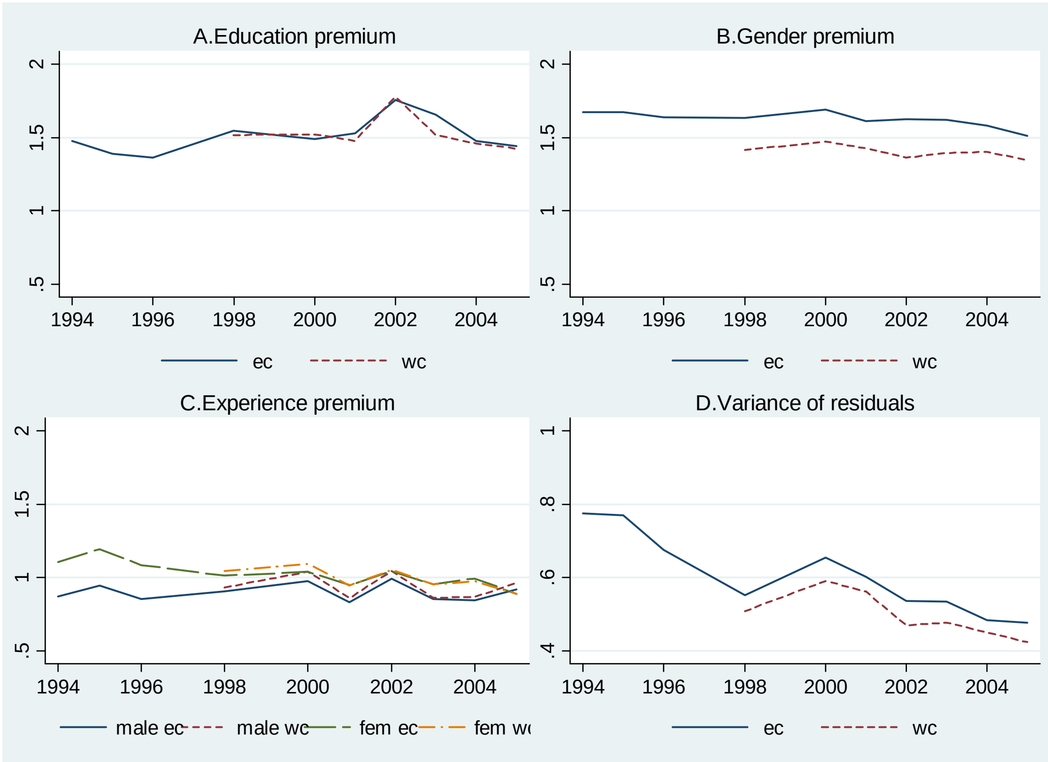
Notes: *wa* = hourly wage rate based on earnings received last month; *wc* = contractual hourly wage rate; *ha* = hours worked last month; *hc* = usual hours of work per month. All wages are deflated with national monthly CPI. Workers are considered full-time if actual hours at primary job were more than 120 hours in the reference month. Panel D compares employment-population ratios in the RLMS sample (R:) and official Goskomstat statistics (G:). Both ratios are calculated for age group 25-59.

Figure 4: Basic Inequality in Individual Wages



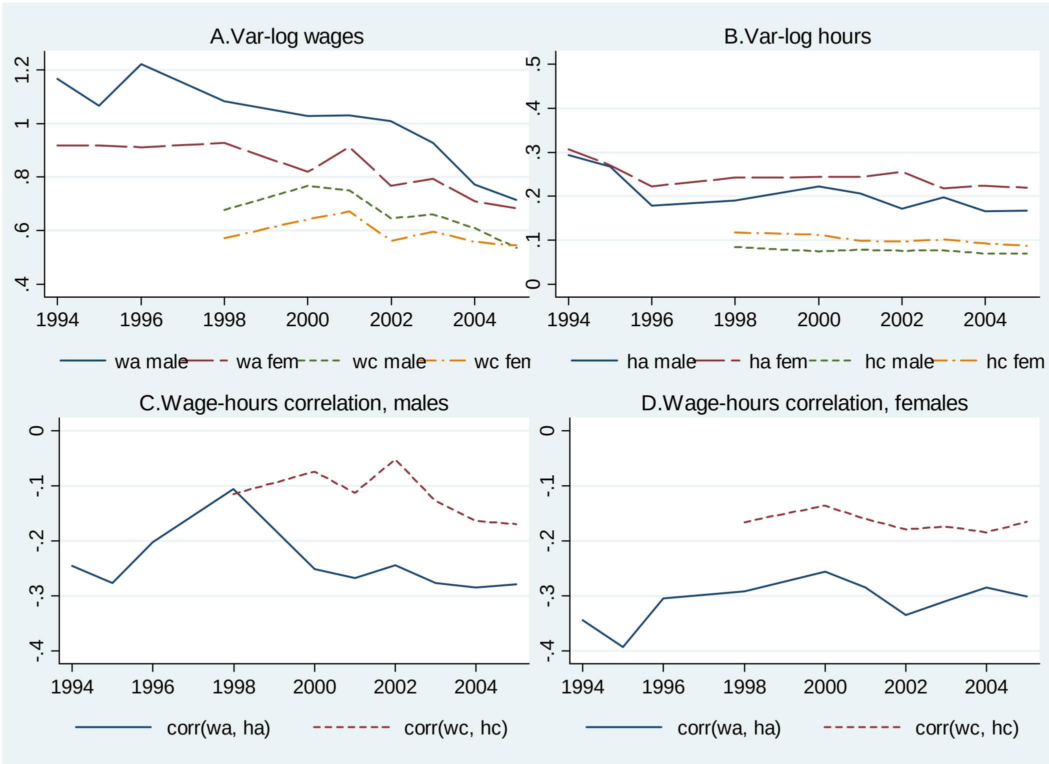
Notes: *ea* = actual individual labor earnings received last month; *ec* = contractual individual labor earnings per month; *wa* = hourly wage rate based on earnings received last month; *wc* = contractual hourly wage rate. Var-log is the variance of log earnings. All earnings are after-tax.

Figure 5: Wage Premia



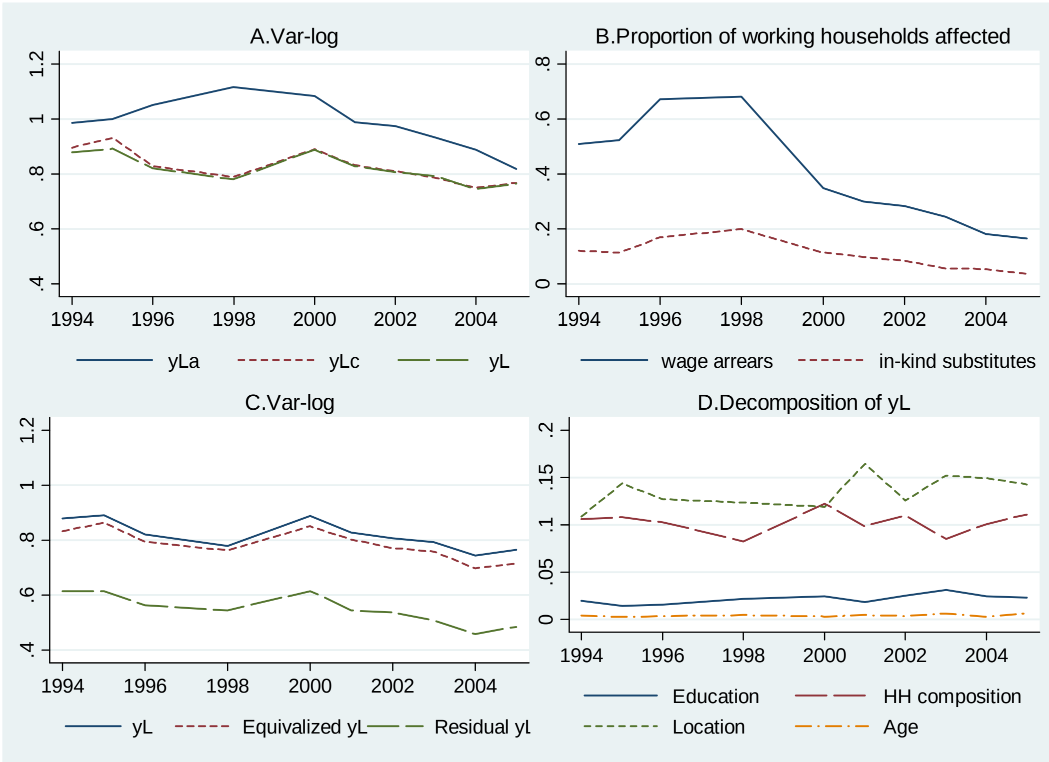
Notes: *ec* = contractual individual labor earnings per month; *wc* = contractual hourly wage rate. All earnings are after-tax. Education premium is the average wage of university educated males divided by the average wage of non university-educated males. Gender premium is the average wage of males divided by the average wage of females. Experience premium is the average wage of age group 45-55 divided by the average wage of age group 25-35. The variance of residuals is from equation (1).

Figure 6: Inequality in Labor Supply



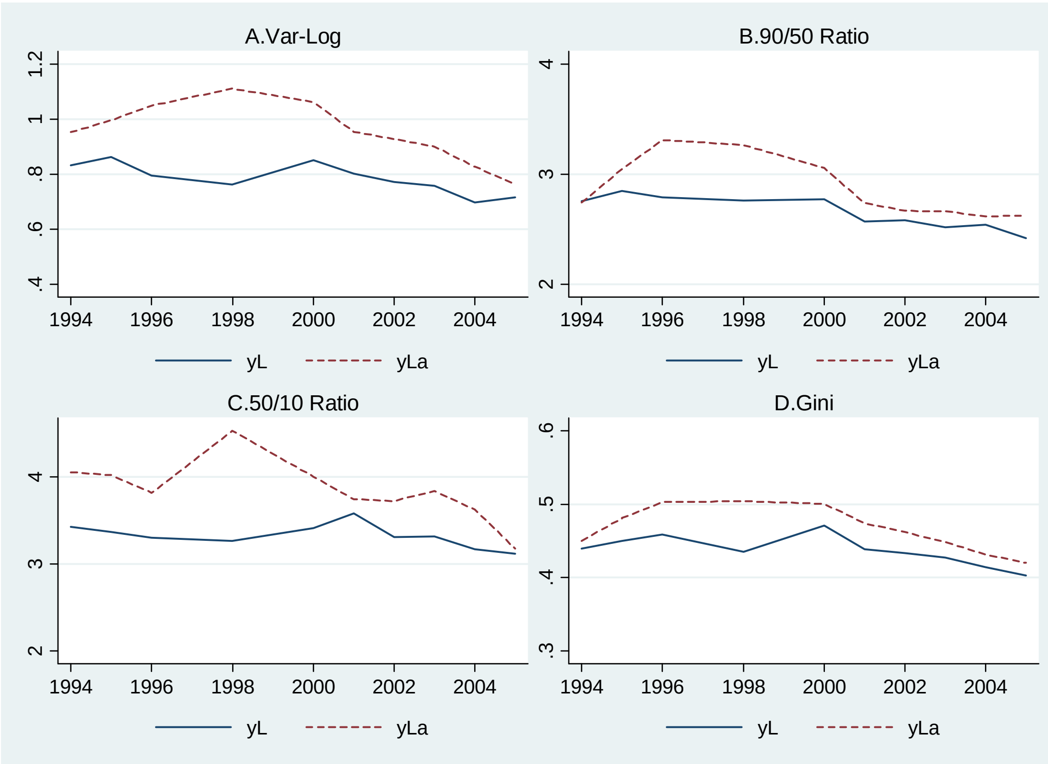
Notes: *wa* = hourly wage rate based on earnings received last month; *wc* = contractual hourly wage rate; *ha* = hours worked last month; *hc* = usual hours of work per month.

Figure 7: Household Earnings Inequality and Its Decomposition



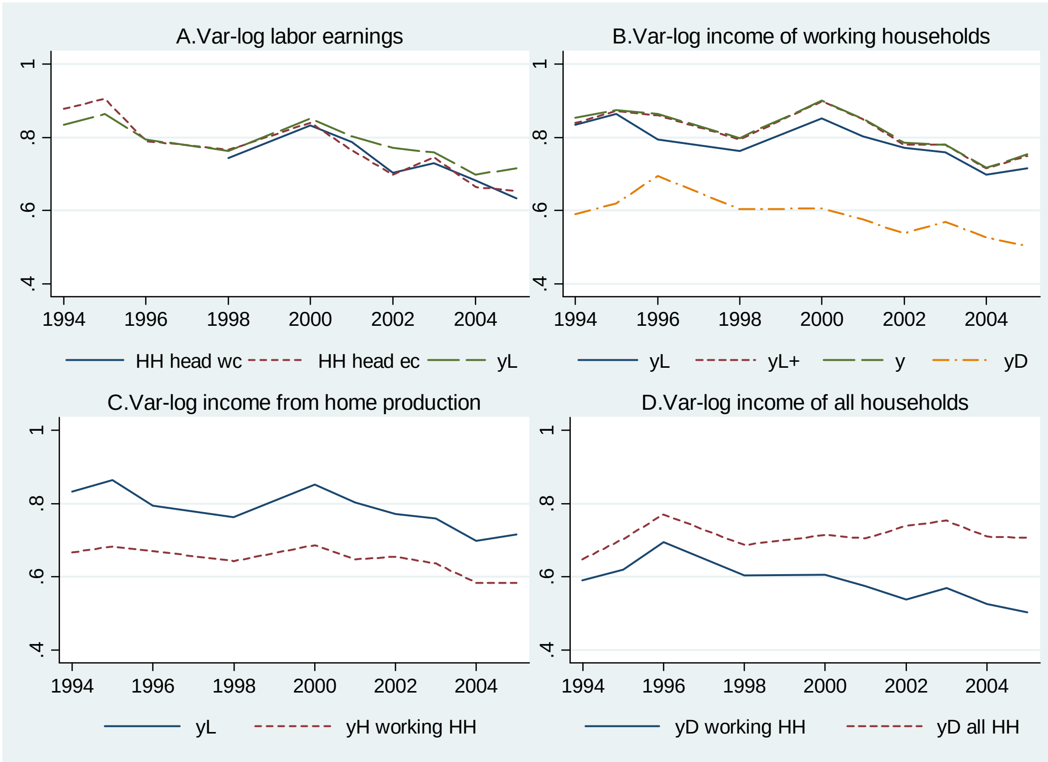
Notes: All earnings are after-tax and deflated using national monthly CPI. *yLa* = actual household labor earnings received last month; *yLc* = household contractual labor earnings per month; *yL* = household contractual labor earnings per month adjusted for non-response. Panel C reports the variance of log raw *yL*, the variance of log *yL* equivalized with an OECD equivalence scale, and the variance of residuals from equation (1). Panel D reports the variance of each observable component of equation (1).

Figure 8: Basic Inequality in Equivalized Household Earnings



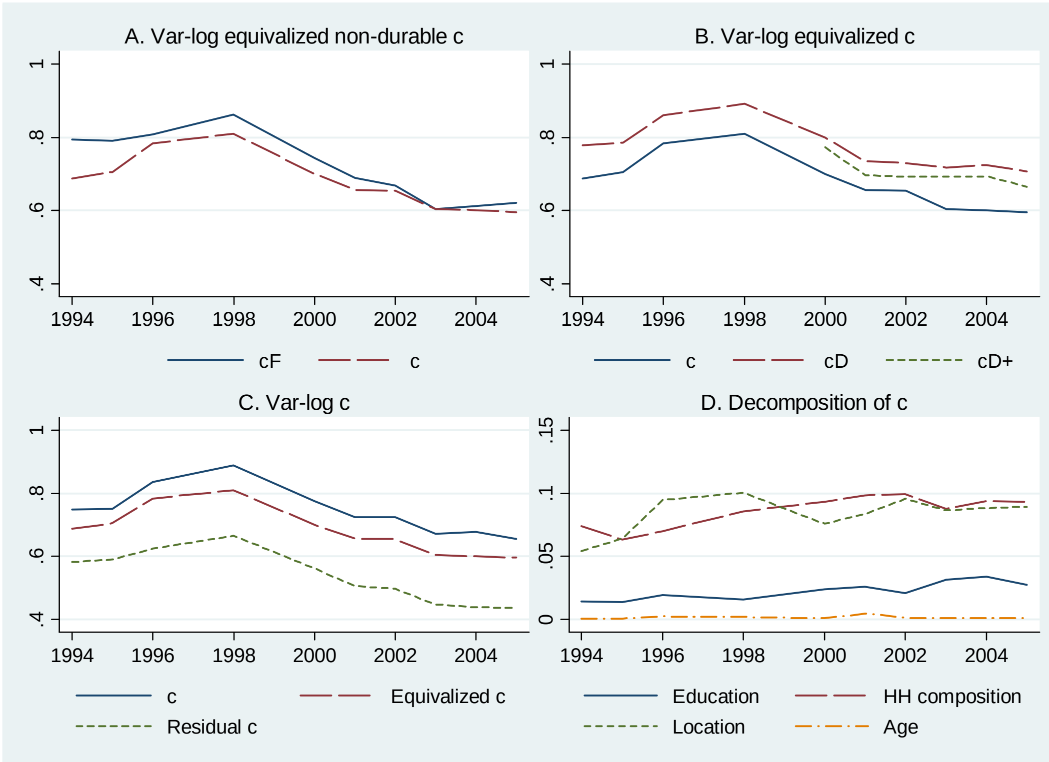
Notes: All earnings are after-tax, equivalized using an OECD equivalence scale, and deflated using national monthly CPI. *yLa* = actual household labor earnings received last month; *yL* = household contractual labor earnings per month adjusted for non-response.

Figure 9: From Wages to Disposable Income



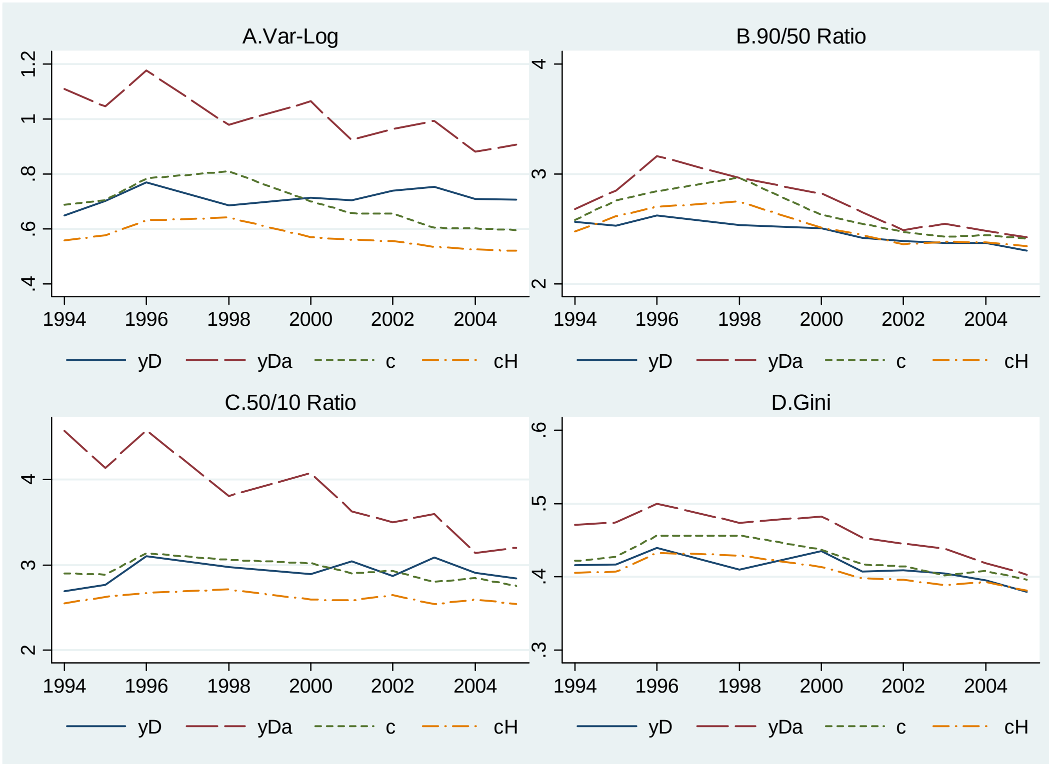
Notes: All income measures are after-tax, equivalized using an OECD equivalence scale, and deflated using national monthly CPI. *HH head wc* = contractual hourly wage rate of the head of household; *HH head ec* = contractual labor earnings per month of the head of household; *yL* = household contractual labor earnings per month adjusted for non-response; *yL+* = *yL+* private transfers; *y* = (*yL+*) + financial income; *yD* = disposable household income = *y* + government transfers; *yH* = *yL* + income from home production. Working households include households with at least one wage earner. Var-log is the variance of the logarithm of income.

Figure 10: Consumption Inequality and Its Decomposition



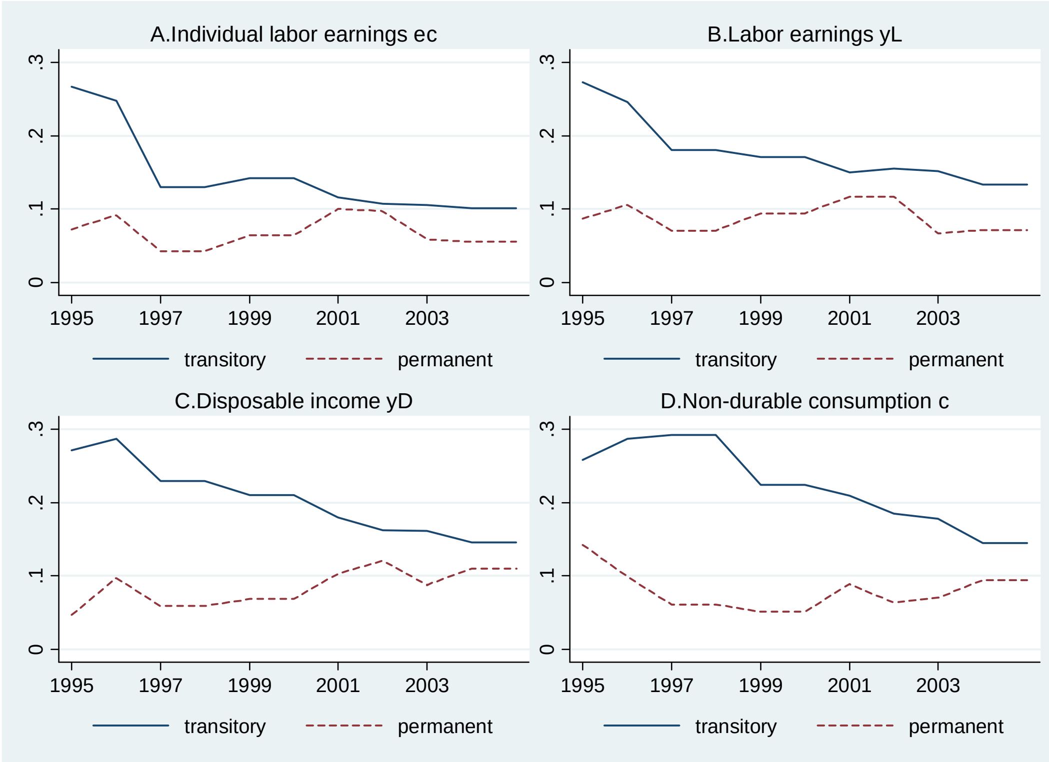
Notes: *cF* = expenditures on food, beverages, and tobacco last week (multiplied by 30/7); *c* = household non-durable expenditures last month; *cD* = *c* + expenditures on durables; *cD+* = *cD* + imputed services from housing. All consumption variables in Panels A and B are per adult equivalent. Panel C reports the variance of log raw *c*, the variance of log *c* equivalized with an OECD equivalence scale, and the variance of the residuals from equation (1). Panel D reports the variance of each observable component from equation (1).

Figure 11: From Disposable Income to Consumption



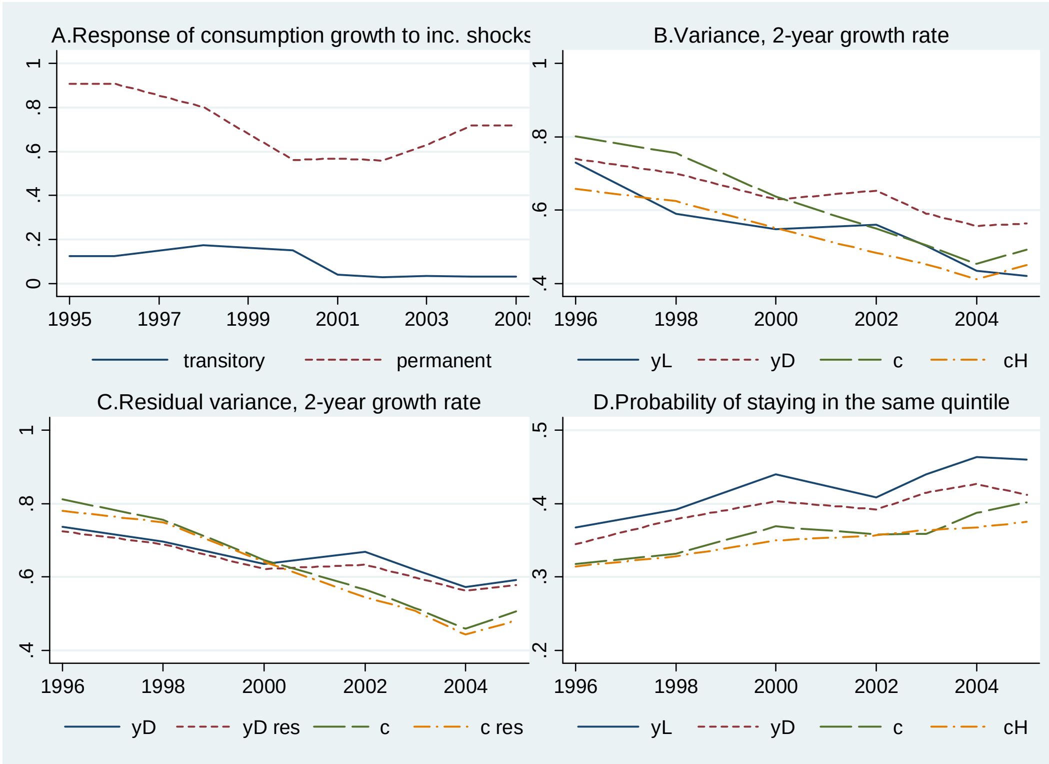
Notes: *yD* = disposable household income based on contractual labor earnings; *yDa* = disposable household income based on actual labor earnings received last month; *c* = household non-durable expenditures last month; *cH* = *c* + consumption of home-grown food. All measures are equivalized using an OECD equivalence scale and deflated with national monthly CPI.

Figure 12: Variance of Shocks in the Estimated Stochastic Processes



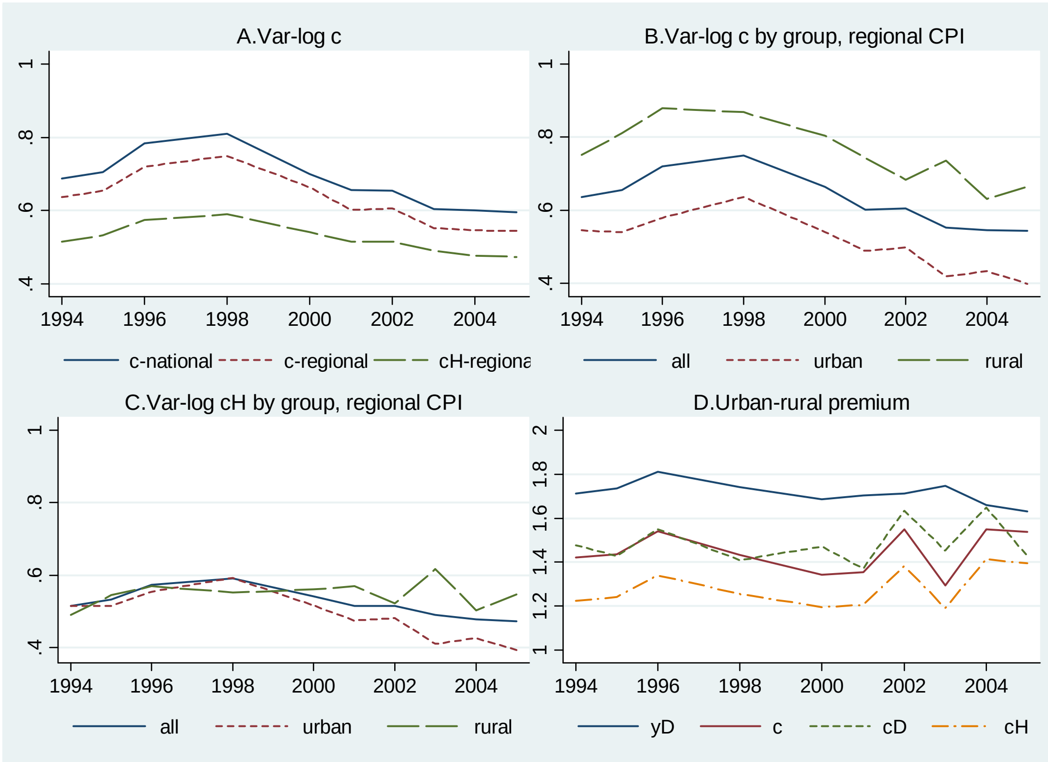
Notes: The figure reports the time series of estimated variance of permanent and transitory shocks. The estimated process is $u_{ht}=\alpha_{ht} + \varepsilon_{ht}$, $\alpha_{ht} = \alpha_{h,t\text{-}1} + \eta_{ht}$, where ε_{ht} is the transitory shock and η_{ht} is the permanent shock. In all specifications, u_{ht} is the residual from projecting the relevant measure of income or consumption on our baseline vector of observable characteristics of households or heads of households. *ec* is contractual labor earnings of household head, *yL* is household contractual labor earnings per month, *yD* is disposable household income based on contractual labor earnings, and *c* is household non-durable expenditures last month. Values in 1998 and 2000 are adjusted for the fact that the permanent shock is accumulated over two years. For both permanent and transitory shocks, 1997 and 1999 values are set equal to 1998 and 2000 values respectively.

Figure 13: Inequality in Income and Consumption Growth



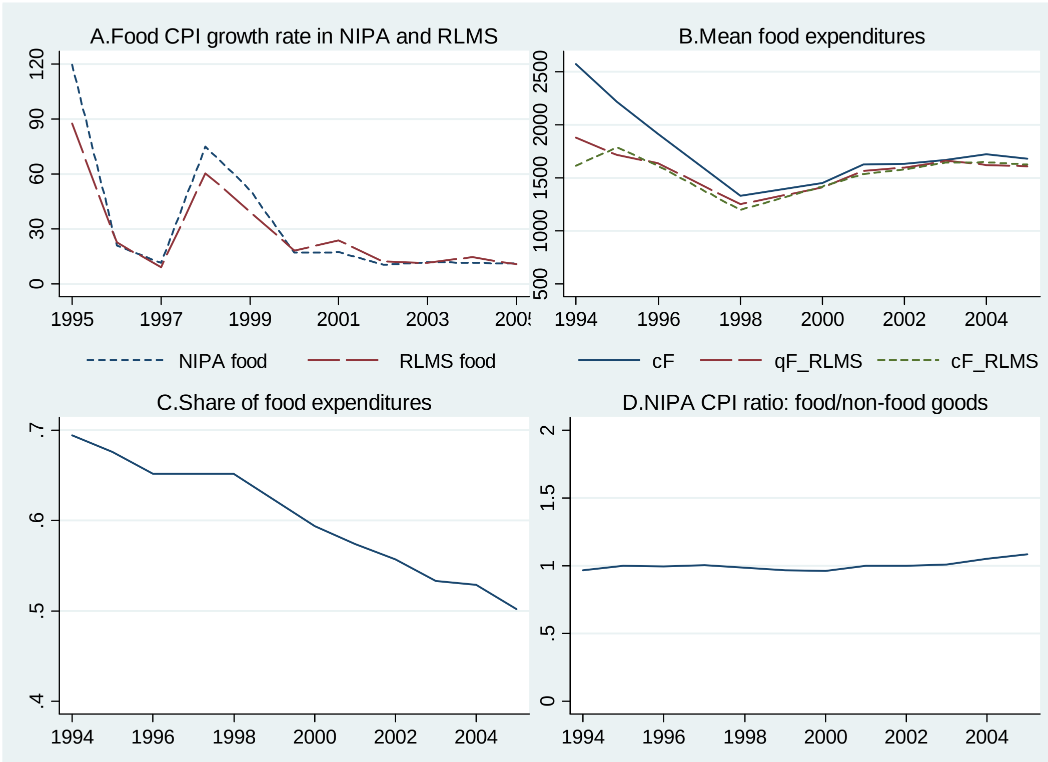
Notes: The responses of consumption to income shocks in panel A are constructed using *yD* for income and *cH* for consumption. See text and notes to Figure 12 for details on constructing shocks. *yL* = household contractual labor earnings per month adjusted for non-response; *yD* = disposable household income based on contractual labor earnings; *c* = household non-durable expenditures last month; *cH* = *c* + consumption of home-grown food. All measures are equivalized using an OECD equivalence scale and deflated with national monthly CPI.

Figure 14: Within-Group and Between-Group Inequality



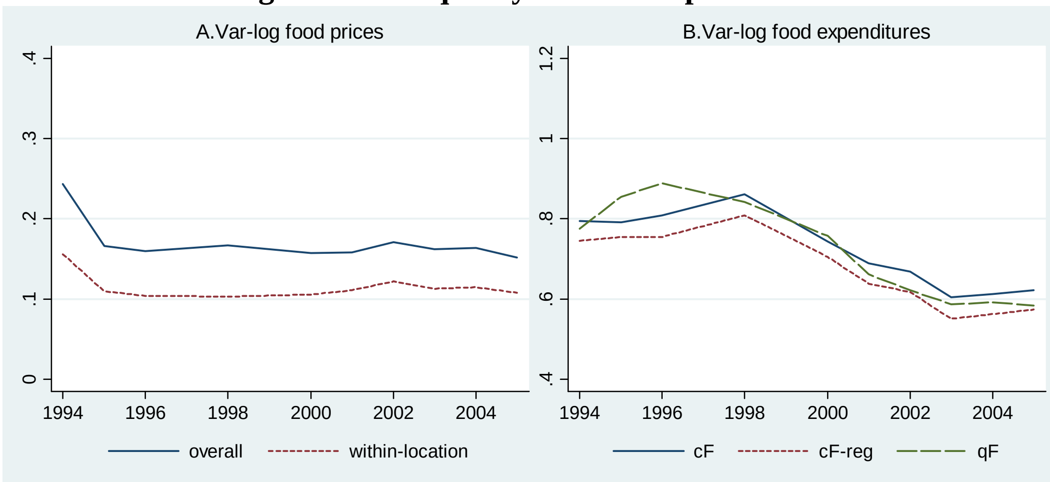
Notes: Rural location is defined as villages and small towns. *yD* = disposable household income based on contractual labor earnings; *c* = household non-durable expenditures last month; *cD* = *c* + expenditures on durables; *cH* = *c* + consumption of home-grown food. All measures are equivalized using an OECD equivalence scale and deflated with regional CPI unless indicated otherwise.

Figure 15: Trends in Food Expenditures



Notes: *cF_RLMS* = expenditures on food, beverages, and tobacco last week (multiplied by 30/7) deflated using national monthly CPI; *cF_RLMS* = expenditures on food, beverages, and tobacco last week deflated using RLMS food CPI; *qF_RLMS* = food quantity index in constant 2002 mean prices for each location. Panel C reports the share of food expenditures *cF* in aggregate consumption expenditures *cD*. All food expenditures are per adult equivalent.

Figure 16: Inequality in Food Expenditures



Notes: *cF* = expenditures on food, beverages, and tobacco last week (multiplied by 30/7) deflated using national monthly CPI; *cF-reg* = *cF* deflated using regional CPI and adjusted for regional differences in cost of living; *qF* = food quantity index in constant 2002 mean prices for each location. All food expenditures are per adult equivalent.

Appendix 1: Data Description

Description of RLMS sample

This study uses ten rounds of the Russian Longitudinal Monitoring Survey (RLMS) that was conducted in 1994-1996, 1998, and 2000-2005. RLMS was not conducted in 1997 and 1999. Time-series reported on the figures are linearly interpolated for missing annual data points.
The RLMS sample consists of the 38 randomly selected primary sample units (municipalities) that are representative of the whole country. They are located in 32 regions (or constituent subjects of the Russian Federation) and 7 federal districts. Russia had 89 constituent subjects and 7 federal districts as of December 1, 2005.

Sample restrictions

We restrict our sample to households in which at least one individual is 25-60 years old. The head of the household in the selected sample is the oldest working-age male or the oldest working-age female if no working-age males are present. If more than one person of the same age-gender is qualified for the head, then the reference person (or the first person surveyed in the roster files) is chosen.

General notes

1. All income variables are *after tax*.

2. All income and consumption variables are constructed on a *monthly basis*.

3. Summary statistics are weighted with individual and household sample weights provided in the RLMS.

4. When a household purchased the item but did not report the amount of the purchase, the missing amounts are imputed by regressing the log of expenditure on the complete interaction between year dummies and federal district dummies, controlling for the size of the household (5 categories), number of children 16 years old or younger (4 categories), number of elderly members 60+ (3 categories), and urban location. Because of the log dependent variable, the predicted values of expenditures are adjusted as $y = exp(\sigma^2/2)exp(\widehat{\log y})$. The subcategories with the largest number of missing values include utilities (2.12% of the sample), gasoline and motor oil (1.63%), transportation services (1.54%), and contributions to non-relatives (1.35%). Missing values for other subcategories are trivial.

5. Similar regression-based imputations are performed for missing subcategories of non-labor income and income from home production. Imputations of labor income are described in the table below. Although the share of missing values for each individual subcategory of non-labor income and expenditures is very small, altogether missing values affect about a third of surveyed households. Our imputation procedure is an improvement over the existing RLMS practice that treats missing values as zeros in computing aggregate income and expenditures.

Variable description and notes

	Variable Name	Definition	Notes
		Individual Earnings and Labor Supply	
ha	Actual hours of work last month	= hours worked last month at primary job + hours worked last month at secondary job + hours spent last month on regular individual economic activities (activities for which an individual is paid for regularly, such as sewing a dress, assisting with repairs, selling goods in a market or on the street, etc.)	Unusually high hours are top coded at 480 hours per month (16 hours per day*30 days)
hc	Usual hours of work per month	= 4 times usual hours in a typical week at primary job + 4 times usual hours in a typical week at secondary job + hours spent last month on regular individual economic activities.	*hc* is available in 1998-2005 only. Unusually high hours are top coded at 480 hours per month (16 hours per day*30 days).
status	Working status	= *full-time* if actual hours at primary job≥120, *part-time* if actual hours at primary job<120, *not working* if a respondent did not work last month at primary job, was not on a temporary leave, and was not engaged in regular individual economic activities	
ea	Actual labor earnings last month	= money received last month from primary job + money received last month from secondary job +	The variable is highly volatile during the period of wage arrears since a worker may

		money received last month from regular individual economic activities + payments in kind received last month from primary job + payments in kind received last month from secondary job	not receive any money last month or receive back payments for several months at once.
ec	Contractual labor earnings per month	*1998-2005, all employees*: = monthly average (over the last 12 months) after-tax labor earnings of an employee at primary job + money received last month from additional jobs for all employees in 1998-2005 *1994-1996, employees with wage arrears*: = total accumulated wage debt divided by the number of months of overdue wages + money received last month from additional jobs for employees with wage arrears at primary job in 1994-1996 *1994-1996, employees with no wage arrears*: = monetary portion of *wa* for employees with no wage arrears *All years, self-employed*: = monetary portion of *wa* for self-employed (or individuals reporting place of work other than an organization), including those involved in regular individual economic activities in all years.	1. *ec* does not include payments in kind. 2. Average monthly earnings are available for an employee at primary job in 1998-2005. 3. Implausibly low earnings below ½ of the official minimum monthly wage are recoded into missing (0.47% of positive earnings). 4. Implausibly high earnings are also recorded into missing if the residuals exceed five standard deviations from the mean after controlling for occupational categories, hours of work, age, age squared, years of schooling, and individual fixed effects (0.13% of positive earnings). 5. For household aggregation purposes, if a respondent worked last month at least one hour but has missing contractual earnings, missing values are imputed using occupational categories, hours of work, gender, age, age squared, years of schooling, urban location and federal district dummies (the share of imputed earnings is 7.8%).
wa	Hourly wage rate last	= *ea* / *ha*	

	month		
wc	Contractual hourly wage rate	= *ec* / *hc*	*hc* is available in 1998-2005 only; *wc* is calculated for non-imputed earnings
		Household Income	
yLa	Actual labor earnings received last month	After-tax payments received by all household members from all places of work in the form of money, goods, and services in the last 30 days as reported by the reference person of the household.	The variable is highly volatile during the period of wage arrears.
yLc	Contractual labor earnings per month	The sum of *ec* across all individual respondents within the household.	Such aggregation omits those adult household members who did not respond to an individual questionnaire; the response rate for working age individuals within the surveyed household is 96.5%.
yL	Contractual labor earnings per month adjusted for non-response	= *yLc* + imputed contractual labor earnings for working-age non-respondents within the household.	Labor earnings of working-age non-respondents are imputed as predicted earnings times the predicted probability of working using the full set of interactions between the four age groups (18-60) and two gender groups and controlling for urban and federal district dummies for each year separately.
yH	Labor earnings plus income from home production	= *yL* + 0.9*h*, where *h* is average monthly income from home-grown food in the last year defined as the sum of physical quantity of produced food items (minus items given away) multiplied by their mean price in a given region, 0.9 is the assumed labor share of home food production.	Mean prices are obtained in two steps. First, the household-specific market price of individual food item is calculated by dividing the cost of purchase by the amount purchased in the last 7 days. Then the mean price of individual food items is computed for each region (*oblast*) and year.

yL+	Labor earnings plus net private transfers	= *yL* + private transfers received last month – private transfers given to individuals outside the household unit last month.	"Private transfers received" include received alimonies and 11 subcategories of contributions from persons outside the household unit, including contributions from relatives, friends, charity, international organizations, etc. "Private transfers given" include alimonies paid and various contributions in money and in kind given to individuals outside the household unit (6 categories).
y	Household income before government transfers	= *yL* + net private transfers + financial income received last month.	Financial income includes dividends on stocks and interest on bank accounts.
yD	Disposable household income	= *y* + public transfers.	Public transfers include government pensions, state child benefits, stipends, unemployment benefits, and government welfare payments.
yDa	Actual disposable household income received last month	= *yLa* + net private transfers + financial income received last month + public transfers.	

Household Consumption

cF	Market expenditures on food, alcohol and tobacco	Monthly expenditures on food, alcohol, and tobacco are computed as the sum of expenditures on individual items in the reference week multiplied by 30/7=4.286.	Items include 50 categories of food at home and away from home, alcoholic and non-alcoholic beverages, and tobacco products. See Appendix 2 for details of computation.
qF	Food quantity index	$qF_t = \sum_k \overline{p}_k q_{kt}$, where q_{kt} is the quantity of food item *k* purchased in year *t* and $\overline{p}_k$ is average price of item *k* for each location (psu) in the base year (2002).	

c	Non-durable expenditures	Sum of expenditures on non-durables in the last 30 days. Non-durable items include food, alcohol, tobacco, clothing and footwear, gasoline and other fuel expenses, rents and utilities, and 15-20 subcategories of services (such as transportation, repair, health care services, education, entertainment, recreation, insurance, etc.).	
cD	Aggregate expenditures	= *c* + expenditures on durables in the last 3 months / 3. Durable items include 10 subcategories such as major appliances, vehicles, furniture, entertainment equipment, etc.	This is compared with purchases of goods and services from NIPA
cH	Non-durable expenditures plus consumption of home-grown food	= *c* + consumption of home-grown food, where the last term is calculated as average monthly quantities of consumed home-grown food items multiplied by their mean price in a given region.	Mean prices are determined in the same way as in *yH*.
cD+	Aggregate expenditures plus services from housing	= *cD* + imputed services from housing.	Imputed services from housing are calculated as 5% of the current housing market value divided by 12.
		Adjustments to Income and Consumption	
equiv	OECD equivalence scale	This equivalence scale assigns a value of 1.0 to the first adult household member, a value of 0.7 to each additional adult, and a value of 0.5 to each child 16 years old and younger.	
cpi_t	National monthly CPI	All income and consumption variables are deflated in prices of 2002 using monthly national CPI.	If the date of interview is in the first half of month, the previous month CPI is used. If the date of interview is in the second half of month, the current month CPI is used.

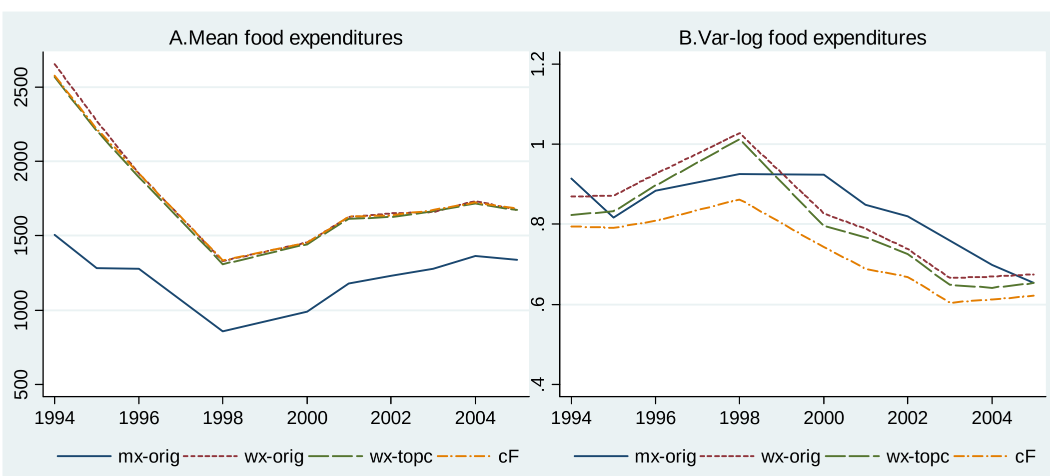
def_t	Regional deflator	Deflator that combines monthly national CPI, December to December regional CPIs, and the regional value of fixed basket of goods and services.	To adjust for monthly inflation, the flow variables are expressed in December prices of each year by using a country average monthly CPI and the date of interview. Next, the annual (December to December) CPI for each 32 oblasts (regions) is applied to convert the flow variables into prices of December 2002. Finally, these real values are adjusted for regional differences in the cost-of-living by using the regional value of fixed basket of goods and services.
$cpiF_{RLMS,t}$	RLMS food CPI	$cpiF_{RLMS,t} = \sum_k p_{kt}\overline{q}_k \Big/ \sum_k \overline{p}_k\overline{q}_k$, where p_{kt} denote the sample average unit price of food category k in year t; $\overline{p}_k$ and $\overline{q}_k$ are the sample average price and the quantity of food item k purchased in the base year.	
		Control Variables	
D^H	Household composition	Vector of household composition variables: 4 categories for the number of children 16 years old and younger (0, 1, 2, and 3+), 3 categories for the number of seniors 60 years old and older (0, 1, and 2+), and 5 categories for the number of household members (1, 2, 3, 4, and 5+).	
	Demographics	A female dummy and continuous age variable a.	
D^E	Schooling	A set of dummies for educational attainment of the head of household (incomplete secondary, secondary, vocational, technical, and university)	
D^L	Location variables	A set of dummies for 7 federal districts, a	

dummy for Moscow and St. Petersburg, and a dummy for urban location.

Appendix 2: Constructing Food Expenditures

This appendix describes the steps in constructing our measure of food expenditures.

1. RLMS food data contain information on the physical quantity and monetary value of last week purchases for 50 categories of food at home and away from home, alcoholic and non-alcoholic beverages, and tobacco products. We first create *wx-orig* as the sum of expenditures on these individual items multiplied by 30/7=4.286. Missing values for this measure are treated as zero.
2. The RLMS questionnaire also asks about the total sum of food purchases in the last 30 days (*mx-orig*). We discard this measure because of a potentially large measurement error, higher probability of underreporting, and ambiguity in the question (e.g., it is likely to exclude beverages and tobacco). We note, however, that the two measures of food expenditures have similar variance (compare *wx-orig* and *mx-orig* in figure below).
3. When a household purchased the item but did not report the quantity of the purchase, the missing quantities are imputed by regressing the log of expenditure on the complete interaction between year dummies and federal district dummies, controlling for the size of the household (5 categories), number of children 16 years old or younger (4 categories), number of elderly members 60+ (3 categories), and urban location. Because of the log dependent variable, the predicted values of expenditures are adjusted as $y = \exp(\sigma^2/2)\exp(\widehat{\log y})$. Missing values for food items are generally trivial.
4. We use top coding of unreasonably high prices in excess of 3 interquantile ranges above the mean prices in a given location as well as unreasonably high amounts (quantities) of food purchases (the top 99th percentile), conditional on the household structure and location. Top coding and imputations does not change the mean value and only slightly reduce the variance (see *wx-topc* in figure below)
5. It is very well known that inequality measures, especially those based on logarithms, are very sensitive to very low values. For that reason, we eliminate the bottom 1% of total food consumption (from purchases and home production) in constant 2002 prices (about 12 percent of the cost of the reference basket of 25 major food items reported by Goskomstat in 2002). While this procedure does not change the mean value of food expenditures, it predictably reduces the variance (see line *cF*).



Notes: All reported measures are per adult equivalent and deflated with national monthly CPI.

Appendix 3: Sample Composition

		Full sample	Restricted sample	Estimation sample
Year:	1994	9.34	9.61	9.66
	1995	8.89	9.09	9.07
	1996	8.82	8.94	8.75
	1998	9.00	9.01	8.91
	2000	9.42	9.24	9.23
	2001	10.64	10.35	10.42
	2002	10.97	10.74	10.81
	2003	11.09	10.92	10.96
	2004	11.07	11.17	11.21
	2005	10.75	10.92	10.99
Region:	Moscow and St. Petersburg	11.28	11.17	11.31
	North West	6.89	7.33	7.37
	Central	19.09	18.17	18.26
	Volga	17.72	17.42	17.39
	South	11.73	12.13	11.93
	Urals	14.17	14.60	14.59
	Siberia	9.41	9.45	9.41
	Far East	9.71	9.73	9.73
Number of household members: 1		18.39	7.58	7.18
	2	27.74	24.28	24.16
	3	25.34	30.83	31.07
	4	18.06	23.49	23.72
	5+	10.47	13.82	13.87
Number of children <16: None		56.99	45.63	45.25
	1	28.26	35.02	35.32
	2	12.23	15.99	16.09
	3+	2.53	3.36	3.34
Urban (excluding small towns)		68.91	69.55	70.01
		42,541	31,969	31,409

Notes: Restricted sample includes households in which at least one individual is 25-60 years old. Estimation sample includes households with non-missing values on disposable income. The sample composition is unweighted.

Appendix 4: Inequality over the Life Cycle

This appendix reports inequality-age regressions. Let $M(a,t)$ denote the cross sectional moment within age group a at time t, such as, for example, the variance of log income among 40-44 year olds in 1994.

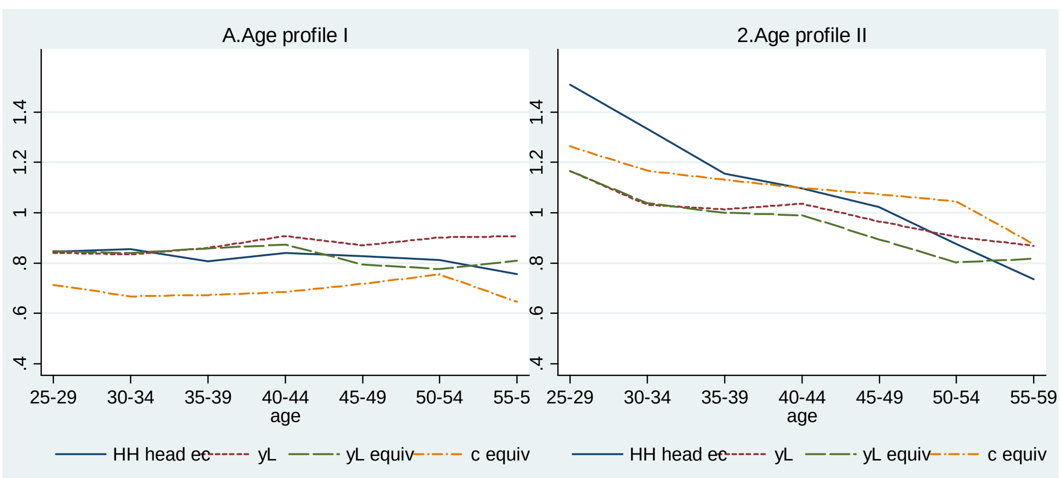
We first perform inequality-age regressions controlling for time effects. We regress the inequality moments on the set of age and time dummies:

$$M(a,t) = \sum_a \beta_a D(a) + \sum_t \beta_t D(t) + \varepsilon_{a,t} .$$

Panel A below shows the pattern of age dummies β_a for different measures of inequality. In almost all cases, the age-inequality profiles are essentially flat, with the exception of a slight decline in inequality among the oldest workers. The flat life cycle inequality profile is consistent with zero contribution of age to inequality decomposition. Panel B reports the age coefficients β'_a from a different specification that assumes away time effects and regresses the cross-sectional inequality moments on age and cohort dummies:

$$M(a,t) = \sum_a \beta'_a D(a) + \sum_{t-a} \beta'_{t-a} D(t-a) + \varepsilon'_{a,t-a} .$$

Now the age-inequality profiles are downward-sloping, because time effects are confounded with age effects. The downward-sloping profile appears to be driven by falling inequality over time.



Notes: Panel A depicts age profiles for the var-log controlling for year effects. Panel B depicts age profiles for the var-log controlling for cohort effects. All measures are deflated with national monthly CPI. *HH head ec* = contractual labor earnings per month of the head of household; *yL* = household contractual labor earnings per month adjusted for non-response; *yL equiv* = *yL* equivalized with an OECD equivalence scale; *c equiv* = household non-durable expenditures equivalized with an OECD equivalence scale.